

83186

Access DB# _____

SEARCH REQUEST FORM

Scientific and Technical Information Center

Requester's Full Name: JANE ZARA Examiner #: 77512 Date: 12/30/02
 Art Unit: 1635 Phone Number 306-5820 Serial Number: 09/915,814
 Mail Box and Bldg/Room Location: 11D03 Results Format Preferred (circle): PAPER DISK E-MAIL
11E12

If more than one search is submitted, please prioritize searches in order of need.

 Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc, if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: AS to H-SL ExpressionInventors (please provide full names): Burton et alEarliest Priority Filing Date: 7/26/01

For Sequence Searches Only Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.

Please Search Seq ID #3

- size unlimited

- size limit to 100 NT's.

Thanks

RECEIVED
 DEC 30 2002
 STIC

NA 3804

Edward Hart
 Technical Info. Specialist
 STIC/Biotech
 CMI 6B02 Tel: 305-9203

STAFF USE ONLY

Type of Search

Vendors and cost where applicable

Searcher: _____	NA Sequence (#) <u>2</u>	STN _____
Searcher Phone #: _____	AA Sequence (#) _____	Dialog _____
Searcher Location: _____	Structure (#) _____	Questel/Orbit _____
Date Searcher Picked Up: <u>12/10/03</u>	Bibliographic _____	Dr.Link _____
Date Completed: <u>1/3/03</u>	Litigation _____	Lexis/Nexis _____
Searcher Prep & Review Time: _____	Fulltext _____	Sequence Systems <u>06</u>
Clerical Prep Time: _____	Patent Family _____	WWW/Internet _____
Online Time: _____	Other _____	Other (specify) _____

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 3, 2003, 03:04:20 ; Search time 6400 Seconds
(without alignments)
17297.980 Million cell updates/sec

Title: US-09-915-814-3
Perfect score: 3804
Sequence: 1 ctcttgtaagagagtgcta.....taataaagtatttaatra 3804

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 2054640 segs, 14551402878 residues
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:
1: gb_da:
2: gb_hlg:
3: gb_in:
4: gb_cm:
5: gb_ov:
6: gb_pat:
7: gb_ph:
8: gb_pl:
9: gb_pr:
10: gb_ro:
11: gb_sts:
12: gb_sy:
13: gb_un:
14: gb_vl:
15: em_da:
16: em_fun:
17: em_hum:
18: em_in:
19: em_mu:
20: em_cm:
21: em_or:
22: em_ov:
23: em_pat:
24: em_ph:
25: em_pl:
26: em_ro:
27: em_sts:
28: em_un:
29: em_vl:
30: em_hlg_hum:
31: em_hlg_inv:
32: em_hlg_other:
33: em_hlg_mus:
34: em_hlg_pln:
35: em_hlg_rod:
36: em_hlg_mam:
37: em_hlg_vit:
38: em_sy:
39: em_higo_hum:
40: em_higo_mus:
41: em_higo_other:

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3804	100.0	3804	6 AX113509	AX113509 Sequence
2	3804	100.0	3804	9 HSU40002	U40002 Human hormo
3	2650.4	69.7	3255	9 HUMSLA	L11706 Human hormo
4	1896.4	49.9	3704	10 RNU40001	U40001 Rattus norv
5	1841.8	48.4	2792	10 AF177401	AF177401 Spermophi
6	1796.6	47.2	3125	4 SSHSLIPAS	AJ000482 Sus scrof
7	1636.4	43.0	3226	10 RATLIP	X51415 Rattus sp.
8	1620	42.6	2678	10 BC021642	BC021642 Mus muscu
9	1619.2	42.6	3172	10 MMU08188	U08188 Mus musculu
10	1155.2	30.4	168586	9 AC011497	AC011497 Homo sapi
11	552.8	14.5	652	4 AY063230	AY063230 Canis fam
12	538.2	14.1	656	4 AF141958	AF141958 Sus scrof
13	531.6	14.0	140808	2 AC021857	AC021857 Homo sapi
14	497.2	13.1	177575	2 AC021857	AJ000483 Sus scrof
15	436	11.5	2222	4 SSHSL5UTR	AJ224692 Sus scrof
16	405.6	10.7	498	4 SSM4692	AF179427 Mus muscu
17	392.8	10.3	17835	10 AF179427	AF179427 Mus muscu
18	392.8	10.3	190860	2 AC073803	AC073803 Mus muscu
19	390.2	10.3	215234	2 AC073745	AC073745 Mus muscu
20	388.6	10.2	131063	2 AC130942	AC130942 Rattus no
21	388.6	10.2	155817	2 AC121639	AC121639 Rattus no
22	282.8	7.4	142394	2 AC109038	AJ006076 Sus scrof
23	272	7.2	42311	4 SSC6076	AC130942 Rattus no
24	269.6	7.1	131063	2 AC130942	GI5903 human STS C
25	258	6.8	258	11 GI5903	AJ237675 Bos tauru
26	256.6	6.7	311	4 BTA237675	AC121639 Rattus no
27	256.4	6.7	155817	2 AC121639	U78042 Bos tauru
28	253	6.7	429	4 BRU78042	AJ060675 Sus scrof
29	245.4	6.5	1513	4 SSC6076	AC004321 Drosophi
30	225.4	5.9	34918	3 AC004321	AC020205 Drosophi
31	225.4	5.9	118533	2 AC020205	AC099023 Drosophi
32	225.4	5.9	180814	3 AC099023	AC099012 Drosophi
33	225.4	5.9	183592	3 AE0039012	AE003792 Drosophi
34	225.4	5.9	256764	3 AE0039012	SE0110 hormone-sen
35	186.6	4.9	225	4 S80110	AC073803 Mus muscu
36	171	4.5	190860	2 AC073803	AY051867 Drosophi
37	149.8	3.9	2328	3 AY051867	AC109038 Rattus no
38	112.8	3.0	142394	2 AC109038	D90903 Synechocyst
39	102	2.7	106209	1 D90903	M17877 P.falciparu
40	99.8	2.6	2359	3 PFARAA	AC0663 P.falciparu
41	96.6	2.5	3187	6 AC0663	AC113928 Rattus no
42	96.4	2.5	137721	2 AC113928	Y17332 Zea mays MR
43	94.6	2.5	1471	8 ZMPROPR	AJ130830 Zea mays
44	94.6	2.5	1792	8 ZMA130830	AY061106 Drosophi
45	87.4	2.3	4186	3 AY061106	

ALIGNMENTS

RESULT 1	AX113509	3804 bp	MRNA	Linear	PAT 01-MAY-2001
LOCUS	AX113509				
DEFINITION	Sequence 1 from Patent WO0126664.				
ACCESSION	AX113509				
VERSION	AX113509.1				
KEYWORDS	GI:13939727				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
TITLE	1 (bases 1 to 3804)				
JOURNAL	Mitchell, G.A. and Wang, S.P.				
	Hormone-sensitive lipase activity mediated male infertility				
	Patent: WO 0126664-A 1 19-APR-2001;				

THIS PAGE BLANK (USPTO)

FEATURES
Source Hospital Sainte-Justine (CA)
Location/Qualifiers
1. 3804
/organism="Homo sapiens"
/db_xref="taxon:9606"
278..3508
/note="unnamed protein product"
/codon_start=1
/protein_id="Cac37855.1"
/db_xref="GI:13939728"
translation="MEPSKSVSRSDMOPEHOPRTPLPEPKEPTPIAOPESKTLQ
SNTQOKPANSNOBRLTOETPAOHDAESOKERAPKQASOEPLAPKAPQOOPYTO
RVLITQOEASQOQPGIGKESITQOEPLARHRAQOPGEPPEPPOQEESEPPAO
AKFGAKSEWAPLESTESQETPEOSDKQTPVQAKSKOGSTELGFTLKQELISQRS
ALRWKALSEWADSESDYSSDSDSPATMGVAGVQSGSTELGFTLKQELISQRS
TSPEHETSAHNRHHYQDPASRLHMDLRTQOGLTVAELNIAFEFSQSGEFAQRG
SGVAFERBOALGLEPALGRILGVALLDEPPEPANGRSGLVNIARCCLAHLHKS
YVANSRSTFERTSHNLAEIAYLAITOLRAIYVYORLITVNPGLPPEEGEGLT
ADPLREYVITLHKCFYGRCLGFOFTPAIRPLOTITIGLVSPGEHYKRNELSVAS
SLTSGRFATIDPELGAEEFRITQNLVHEWKAEMWNTTEMEVLSLAMASATVRSR
LSLPEAFEMPLTADPTLVITISPLAHGPGPVLYRLISYDLREGDSELSLIK
SNGORSLELWPPQOAPRSRLIVHFGGGEVQAOISHEBYLKSMAOELAPITSID
YSLAPEAPFRALEECEFAVCMAIKHCAIGSTGERICLADSGAGCFVALRAAA
YGYRVDGIMAAVPAATMLOPASPSRLISMDPLIPLSKYCSATYAGATEDHSNS
DQKALGMGLVRDRLALLDFRLGASSWLSLELSGRKSQKASEPLAEPMRSVSE
AALAQPGPLGTDLSLKNLTLRLDLSLGNSETSDTPKMSLAETLSPTGVDNPLP
PEDAGEAEAKNELSPMDRGIVYAAAPFEGHPRRSOGATPESPLIKNFPMSP
LLAPDSMLKSLPVHIVACALDPLDSDVMILRLNIGOPVTLRAVEDLPHGLTLA
ALCREFTOAELEVERIRLVLTLPAGAPSEGTGAGVDSGCGGRH"
BASE COUNT 828 a 1225 c 1091 g 660 t
ORIGIN
Query Match 100.0%; Score 3804; DB 6; Length 3804;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3804; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CTTCTTTAAGAGAGTGTAGGCACATAGCCCCCTCTATTCTTATCTCCACCACCAAG 60
DB 1 CTTCTTTAAGAGAGTGTAGGCACATAGCCCCCTCTATTCTTATCTCCACCACCAAG 60
QY 61 AAGAGGACAGAGTTCCTACTTAGTGGGGCCAGCTGTGATCGGCCAACAGCCAGCTG 120
DB 61 AAGAGGACAGAGTTCCTACTTAGTGGGGCCAGCTGTGATCGGCCAACAGCCAGCTG 120
QY 121 CTTTAAAAAGAGACCAAGTATGCTAGATGAGTGAACCCCAAGAGAACTGCCATCA 180
DB 121 CTTTAAAAAGAGACCAAGTATGCTAGATGAGTGAACCCCAAGAGAACTGCCATCA 180
QY 181 TGAGGAATCATGAGATCTGTGAAGAGAGAGGCTGGTGGAGCCCAAGAGATAGA 240
DB 181 TGAGGAATCATGAGATCTGTGAAGAGAGAGGCTGGTGGAGCCCAAGAGATAGA 240
QY 241 ACCTGAAGATCAATATCTCCGTGAGGGAATAACATGAGCCAGGTTCTTAAGTCACT 300
DB 241 ACCTGAAGATCAATATCTCCGTGAGGGAATAACATGAGCCAGGTTCTTAAGTCACT 300
QY 301 GTCTAGGTACAGTGGCAACTGGAACACACAGAGGCTATTAACCCCGCTAGAGCCCTG 360
DB 301 GTCTAGGTACAGTGGCAACTGGAACACACAGAGGCTATTAACCCCGCTAGAGCCCTG 360
QY 361 GCCGAAAAAGACCCATAGGCCAGCCAGCAATGCAAGACTCGCAGGGATCCAAATACCA 420
DB 361 GCCGAAAAAGACCCATAGGCCAGCCAGCAATGCAAGACTCGCAGGGATCCAAATACCA 420
QY 421 ACAGAAAGCTGCTTCAAAACCAAGACCCCTCACCAGAGAGAGACCCCTGCACAACATGA 480
DB 421 ACAGAAAGCTGCTTCAAAACCAAGACCCCTCACCAGAGAGAGACCCCTGCACAACATGA 480
QY 481 TGTGTAATCCAGAGAACTAGAGCCCAACAAAATCTGCTTCAACAAGAGAAATTTCT 540
DB 481 TGTGTAATCCAGAGAACTAGAGCCCAACAAAATCTGCTTCAACAAGAGAAATTTCT 540
QY 541 TGGCCCAAGAGAGCCGACACAGCAATCACTTACATCCAAAAGGTGCTCTACATCA 600

DB 541 TGGCCCAAGAGAGCCGACACAGCAATCACTTACATCCAAAAGGTGCTCTACATCA 600
QY 601 ACAGAGAGTGGCTCCACAGAGGAGACCTGGGTAGGAAAAATCTATTAACATAAGAGA 660
DB 601 ACAGAGAGTGGCTCCACAGAGGAGACCTGGGTAGGAAAAATCTATTAACATAAGAGA 660
QY 661 GCCAGCATTAGACAAAAGCATGTAGCCAGCCAGGCGCTGGGCCAGAGAGCCACTTCC 720
DB 661 GCCAGCATTAGACAAAAGCATGTAGCCAGCCAGGCGCTGGGCCAGAGAGCCACTTCC 720
QY 721 ACCTCAACAAAGAGCTGATCAACACCTGGGCCAGGCTTAACCTGGAGCCAAAAGGA 780
DB 721 ACCTCAACAAAGAGCTGATCAACACCTGGGCCAGGCTTAACCTGGAGCCAAAAGGA 780
QY 781 GCCATCTGCCCGACTGATCTACATCCCAAGAGACCTGGAACAGTCAAGACAAGAAC 840
DB 781 GCCATCTGCCCGACTGATCTACATCCCAAGAGACCTGGAACAGTCAAGACAAGAAC 840
QY 841 AACGCCAGTCCAGGAGGCCAAATCCAAAGAGGAGATCTTGACAGAGCTGGGATTTCTAAC 900
DB 841 AACGCCAGTCCAGGAGGCCAAATCCAAAGAGGAGATCTTGACAGAGCTGGGATTTCTAAC 900
QY 901 AAAACTTCAGGAATCTCATACAGGATCAGCCCTAGAGTGGAGAGCACTTCTAGTG 960
DB 901 AAAACTTCAGGAATCTCATACAGGATCAGCCCTAGAGTGGAGAGCACTTCTAGTG 960
QY 961 GGTGCGAGATCTGAGTCAAGATCAAGATGGGATATATCTTCAGACACAGATCTCCAGC 1020
DB 961 GGTGCGAGATCTGAGTCAAGATCAAGATGGGATATATCTTCAGACACAGATCTCCAGC 1020
QY 1021 CACGATGGGTGAATGTGGGCCAGGAGTGAAGCTAGGCTTCAAGAAAAATCTGGTTA 1080
DB 1021 CACGATGGGTGAATGTGGGCCAGGAGTGAAGCTAGGCTTCAAGAAAAATCTGGTTA 1080
QY 1081 TAAAGTATGTAGATTAAGTGGGACGTCCGACATGAGAAAACAGTCTCGGAATCA 1140
DB 1081 TAAAGTATGTAGATTAAGTGGGACGTCCGACATGAGAAAACAGTCTCGGAATCA 1140
QY 1141 CAGACACTACAGAGATACAGCCCTCAAGGCTCATCCAAATGAGACTGGCAATAGAC 1200
DB 1141 CAGACACTACAGAGATACAGCCCTCAAGGCTCATCCAAATGAGACTGGCAATAGAC 1200
QY 1201 ACAGTGGTGTGACTCTGGCGGAGGACAATAGCCTTCTCTCAGCCAGGCTCTG 1260
DB 1201 ACAGTGGTGTGACTCTGGCGGAGGACAATAGCCTTCTCTCAGCCAGGCTCTG 1260
QY 1261 GGAAGAGGCCAGGCGCTGACAGGCTTTTGGCGGTGACAGGAGACAGGCGCTGGGCT 1320
DB 1261 GGAAGAGGCCAGGCGCTGACAGGCTTTTGGCGGTGACAGGAGACAGGCGCTGGGCT 1320
QY 1321 GGAGCGGCGCTGGGCGGCTGCTGGGTGGGCACTTTGAGACCTGGACCCAGAGAC 1380
DB 1321 GGAGCGGCGCTGGGCGGCTGCTGGGTGGGCACTTTGAGACCTGGACCCAGAGAC 1380
QY 1381 ACCGCGCAACGGGTACCGAGACCTAGTGAACACAGCCCGCTGCTGGCGCACTCTCT 1440
DB 1381 ACCGCGCAACGGGTACCGAGACCTAGTGAACACAGCCCGCTGCTGGCGCACTCTCT 1440
QY 1441 GCACAAATCCGCTATGTGGGCTCCAAACGCGGACACATCTTCCGACCAAGCCACAA 1500
DB 1441 GCACAAATCCGCTATGTGGCTCCAAACGCGGACACATCTTCCGACCAAGCCACAA 1500
QY 1501 CTTGGCGAGTGAAGGCTTACCTGAGCTCCCTCACCAGACTCCGCGCTGGTCTACTA 1560
DB 1501 CTTGGCGAGTGAAGGCTTACCTGAGCTCCCTCACCAGACTCCGCGCTGGTCTACTA 1560
QY 1561 CGCCCAAGCGCTGTGTGTTACCAATGCGCGGGGTACTCTTCTTGAAGGCGACAGAGG 1620
DB 1561 CGCCCAAGCGCTGTGTGTTACCAATGCGCGGGGTACTCTTCTTGAAGGCGACAGAGG 1620
QY 1621 GGTCAACCGCGACTCTGCTGGGAGATGTCAGCGTGAAGAGATGCTTCAATGGCG 1680
DB 1621 GGTCAACCGCGACTCTGCTGGGAGATGTCAGCGTGAAGAGATGCTTCAATGGCG 1680

THIS PAGE BLANK (USPTO)

Db 1621 GCTCACCGCCGACTTCTCCGGAGTATGTCACGCTGCATAGAGGATGCTTATGCGCC 1680
QY 1681 CTGCTGGGCTTCAGTTCAAGCTGCTGCATCCGGCATTTCTGAGACATCTCATTTG 1740
Db 1681 CTGCTGGGCTTCAGTTCAAGCTGCTGCATCCGGCATTTCTGAGACATCTCATTTG 1740
QY 1741 GCTGCTCTCTTGGGGGAGCACTACAAAGCAGAGAGAGGCTCAGTGTGGCCGAG 1800
Db 1741 GCTGCTCTCTTGGGGGAGCACTACAAAGCAGAGAGAGGCTCAGTGTGGCCGAG 1800
QY 1801 CTCTCTCTTACAGAGGCGGCTTGGCATGACCCGAGCTGGTGGGCTGAGTTGA 1860
Db 1801 CTCTCTCTTACAGAGGCGGCTTGGCATGACCCGAGCTGGTGGGCTGAGTTGA 1860
QY 1861 GCGGATACACAGAACCTTGAGCTGACCTTCTGAAAAGCTTCTGAAACATCACGAGAT 1920
Db 1861 GCGGATACACAGAACCTTGAGCTGACCTTCTGAAAAGCTTCTGAAACATCACGAGAT 1920
QY 1921 GGAAGTCTATGCTCTTGGGCAACATGAGCATGCGGCAAGGTAGAGGCTGCT 1980
Db 1921 GGAAGTCTATGCTCTTGGGCAACATGAGCATGCGGCAAGGTAGAGGCTGCT 1980
QY 1981 CACCTGACACCCGAAAGCTTTGAGATGCCACTGACGAGCCGACGCTCAAGGTAC 2040
Db 1981 CACCTGACACCCGAAAGCTTTGAGATGCCACTGACGAGCCGACGCTCAAGGTAC 2040
QY 2041 CATCTCACCCCACTGGCCACACAGGCTTGCGCCGCTCTGCTGAGGCTCATCTCTA 2100
Db 2041 CATCTCACCCCACTGGCCACACAGGCTTGCGCCGCTCTGCTGAGGCTCATCTCTA 2100
QY 2101 TGACCTGCTGGAAGAGACAGACAGTGAAGCTGACAGCTGATTAAGTCCAAAGGCA 2160
Db 2101 TGACCTGCTGGAAGAGACAGACAGTGAAGCTGACAGCTGATTAAGTCCAAAGGCA 2160
QY 2161 ACAGAGCTGAGAGCTGTGGCGGCGCCGACAGAGCAAGCCGCTGGGCTCCGATAGT 2220
Db 2161 ACAGAGCTGAGAGCTGTGGCGGCGCCGACAGAGCAAGCCGCTGGGCTCCGATAGT 2220
QY 2221 GCACCTTCACAGGCGGTGGCTTTGTGGCCAGACCTCCAGATCCAGAGGCTCACTCAA 2280
Db 2221 GCACCTTCACAGGCGGTGGCTTTGTGGCCAGACCTCCAGATCCAGAGGCTCACTCAA 2280
QY 2281 GAGCTGGGCGCCAGAGAGCTGGGGCGCCCATCATCTCCATGACATCTCCCTGAGCCCTGA 2340
Db 2281 GAGCTGGGCGCCAGAGAGCTGGGGCGCCCATCATCTCCATGACATCTCCCTGAGCCCTGA 2340
QY 2341 GGGCCCTTCCCGCGTGGCGGTGGAGAGTGTCTTGGCTACAGCTGGGGCATCAAGCA 2400
Db 2341 GGGCCCTTCCCGCGTGGCGGTGGAGAGTGTCTTGGCTACAGCTGGGGCATCAAGCA 2400
QY 2401 CTGCGGCTCTCTTGGCTCAACAGAGGGAACGATCTGCTTGGCGGGGAGAGTGCAGGCG 2460
Db 2401 CTGCGGCTCTCTTGGCTCAACAGAGGGAACGATCTGCTTGGCGGGGAGAGTGCAGGCG 2460
QY 2461 GAACTCTCTCTTCAAGCTGTGCTCTTGGGAGAGAGCTTACGAGGAGTGGGGTGCAGATGG 2520
Db 2461 GAACTCTCTCTTCAAGCTGTGCTCTTGGGAGAGAGCTTACGAGGAGTGGGGTGCAGATGG 2520
QY 2521 CATCATGAGAGCTTACCGGCGCAATGCTGAGGCTGGCGGCTCTCCGCGCGCT 2580
Db 2521 CATCATGAGAGCTTACCGGCGCAATGCTGAGGCTGGCGGCTCTCCGCGCGCT 2580
QY 2581 GAGGCTCATGAGAGCTTGGCTGCGGCTCACTGCTGCTTCCAAAGTGTGACAGGCTATGC 2640
Db 2581 GAGGCTCATGAGAGCTTGGCTGCGGCTCACTGCTGCTTCCAAAGTGTGACAGGCTATGC 2640
QY 2641 TGGTGAAGAAGAGAGAGACATCCAACTAGACAGACAAAAGCTCGGCAATGATGGGCT 2700
Db 2641 TGGTGAAGAAGAGAGAGACATCCAACTAGACAGACAAAAGCTCGGCAATGATGGGCT 2700
QY 2701 GGTGCGGCGGAGACAGAGCTGCTCCGAGAGCTTCCGCTGGGCTGCTCATGGCT 2760
Db 2701 GGTGCGGCGGAGACAGAGCTGCTCCGAGAGCTTCCGCTGGGCTGCTCATGGCT 2760

QY 2761 CAACTCTCTCTGAGTTAAGTGGGCGCAAGTCCAGAGAGATGTGAGACCCATAGCAGA 2820
Db 2761 CAACTCTCTCTGAGTTAAGTGGGCGCAAGTCCAGAGAGATGTGAGACCCATAGCAGA 2820
QY 2821 GCGGATGCGCGCAGTGTGTGAAGACAGACTGGCCAGGCCAGGCGCCACTGGGAC 2880
Db 2821 GCGGATGCGCGCAGTGTGTGAAGACAGACTGGCCAGGCCAGGCGCCACTGGGAC 2880
QY 2881 GGAATTCCTCAAGAACTGAGGCTGAGGAGCTTGAAGGCTGAGGAGAACTCCGAGACTG 2940
Db 2881 GGAATTCCTCAAGAACTGAGGCTGAGGAGCTTGAAGGCTGAGGAGAACTCCGAGACTG 2940
QY 2941 GTGAGACACCCCGAGATGTGCTGCTGAGAGACTGAGACTTGGCCCTCACACCTCCGA 3000
Db 2941 GTGAGACACCCCGAGATGTGCTGCTGAGAGACTGAGACTTGGCCCTCACACCTCCGA 3000
QY 3001 TGTCACTCTTATTTACCACTGAGAGATGACAGGGAGAGGCTGAGGCCAAAATGAGCT 3060
Db 3001 TGTCACTCTTATTTACCACTGAGAGATGACAGGGAGAGGCTGAGGCCAAAATGAGCT 3060
QY 3061 GAGCCCATGAGACAGAGGCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3120
Db 3061 GAGCCCATGAGACAGAGGCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3120
QY 3121 AAGCTCCAGCGAGGTGCGACACAGATGCGCTTACTCTACACCATAGTCAAGAACCC 3180
Db 3121 AAGCTCCAGCGAGGTGCGACACAGATGCGCTTACTCTACACCATAGTCAAGAACCC 3180
QY 3181 CTTCATGTGCGCGCTGCTGCGACCCGACAGAGATGCTCAAGAGCTGACCTGTGCAT 3240
Db 3181 CTTCATGTGCGCGCTGCTGCGACCCGACAGAGATGCTCAAGAGCTGACCTGTGCAT 3240
QY 3241 CGTGCGGCGGCGCTGAGACCCCATGCTGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3300
Db 3241 CGTGCGGCGGCGCTGAGACCCCATGCTGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3300
QY 3301 CAACCTGGGCGAGCGGTGAGAGCTGCGGCTGAGAGAGCTGCGGCGAGCTTCTGAG 3360
Db 3301 CAACCTGGGCGAGCGGTGAGAGCTGCGGCTGAGAGAGCTGCGGCGAGCTTCTGAG 3360
QY 3361 CCTAGCGGCGGTGTGCGCGAGAGCGCGCAGAGCGCAGAGCTGTGCTGAGAGCGCATCG 3420
Db 3361 CCTAGCGGCGGTGTGCGCGAGAGCGCGCAGAGCGCAGAGCTGTGCTGAGAGCGCATCG 3420
QY 3421 CCTGCTCTCACTCTCTCCCGCGAGAGCGGCGCGAGCGGAGAGAGGCGGCTGCGGGGT 3480
Db 3421 CCTGCTCTCACTCTCTCCCGCGAGAGCGGCGCGAGCGGAGAGAGGCGGCTGCGGGGT 3480
QY 3481 AGAGGCGGCTGCGGCGGAGCACTAAAGCTGTGTTGCCATCTGCGCGGCTCG 3540
Db 3481 AGAGGCGGCTGCGGCGGAGCACTAAAGCTGTGTTGCCATCTGCGCGGCTCG 3540
QY 3541 TCATGAATGCTTCCGGGCGCGGCGGAGAGGAGCGGCGGCTGTGCTTAAAGTGGGG 3600
Db 3541 TCATGAATGCTTCCGGGCGCGGCGGAGAGGAGCGGCGGCTGTGCTTAAAGTGGGG 3600
QY 3601 GTGGCAAGGCGGCGGCGGCGGCGGAGAGCTGAGAGCTGCGGAGAGGAGGAGGAGAG 3660
Db 3601 GTGGCAAGGCGGCGGCGGCGGCGGAGAGCTGAGAGCTGCGGAGAGGAGGAGGAGAG 3660
QY 3661 GCACACACAGCGGTACAGAGAGAGGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAG 3720
Db 3661 GCACACACAGCGGTACAGAGAGAGGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAG 3720
QY 3721 CTGCTGCGGCGAGCGCGGAGAGAGGAGAGGAGCTGCGCTCTGCGAGTGTGGGTTT 3780
Db 3721 CTGCTGCGGCGAGCGCGGAGAGAGGAGAGGAGAGGAGCTGCGCTCTGCGAGTGTGGGTTT 3780
QY 3781 GTTGAATTAAGATTTAATTA 3804
Db 3781 GTTGAATTAAGATTTAATTA 3804

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 3, 2003, 03:04:20 : Search time 6400 Seconds
(without alignments)
17297.980 Million cell updates/sec

Title: US-09-915-814-3

Perfect score: 3804

Sequence: 1 cttcttgtaagagagtgtcta.....taataaaagtatttaatta 3804

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*

1: gb_ba:*

2: gb_hlg:*

3: gb_in:*

4: gb_om:*

5: gb_ov:*

6: gb_pat:*

7: gb_ph:*

8: gb_pl:*

9: gb_pr:*

10: gb_ro:*

11: gb_sts:*

12: gb_sy:*

13: gb_un:*

14: gb_vl:*

15: em_ba:*

16: em_fun:*

17: em_hum:*

18: em_in:*

19: em_mu:*

20: em_om:*

21: em_or:*

22: em_ov:*

23: em_pat:*

24: em_ph:*

25: em_pl:*

26: em_ro:*

27: em_sts:*

28: em_un:*

29: em_vl:*

30: em_hlg_hum:*

31: em_hlg_inv:*

32: em_hlg_other:*

33: em_hlg_mus:*

34: em_hlg_pln:*

35: em_hlg_rtd:*

36: em_hlg_mam:*

37: em_hlg_vrt:*

38: em_sy:*

39: em_hlgo_hum:*

40: em_hlgo_mus:*

41: em_hlgo_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	length	DB ID	Description
1	3804	100.0	3804	6 AX113509	AX113509 Sequence
2	3804	100.0	3804	9 HSU04002	U40002 Human hormo
3	2650.4	69.7	3255	9 HUMSLA	L11706 Human hormo
4	1896.4	49.9	3704	10 RNU40001	U40001 Rattus norv
5	1841.8	48.4	2792	10 AF177401	AF177401 Spermoph
6	1796.6	47.2	3125	4 SSHSLIPS	AJ000482 Sus scrof
7	1636.4	43.0	3226	10 RAITIP	X51415 Rattus sp.
8	1620	42.6	2678	10 BC021642	BC021642 Mus muscu
9	1619.2	42.6	3172	10 MM008188	U08188 Mus musculu
10	1155.2	30.4	168586	9 AC011497	AC011497 Homo sapi
11	552.8	14.5	652	4 AY063230	AY063230 Canis fam
12	538.2	14.1	656	4 AF141958	AF141958 Sus scrof
13	531.6	14.0	140808	2 AC021858	AC021858 Homo sapi
14	497.2	13.1	177575	2 AC021857	AC021857 Homo sapi
15	436	11.5	2222	4 SSHSL50TR	AJ224692 Sus scrof
16	405.6	10.7	498	4 SSAJ4692	AF179427 Mus muscu
17	392.8	10.3	17835	10 AC073803	AC073803 Mus muscu
18	392.8	10.3	190860	2 AC073745	AC073745 Mus muscu
19	390.2	10.3	215234	2 AC130942	AC130942 Rattus no
20	388.6	10.2	131063	2 AC121639	AC121639 Rattus no
21	388.6	10.2	155817	2 AC109038	AC109038 Rattus no
22	282.8	7.4	142394	2 SSC6076	AJ006076 Sus scrof
23	272	7.2	4911	4 AC130942	AC130942 Rattus no
24	269.6	7.1	131063	2 G15903	G15903 human STS C
25	258	6.8	258	11 BTA237675	AJ237675 Bos tauru
26	256.6	6.7	311	4 AC121639	AC121639 Rattus no
27	256.4	6.7	155817	4 BTU78042	U78042 Bos taurus
28	253	6.7	429	4 SSC6075	AJ006075 Sus scrof
29	245.4	6.5	1513	4 AC004321	AC004321 Drosophi
30	225.4	5.9	34918	3 AC020205	AC020205 Drosophi
31	225.4	5.9	118533	2 AC099023	AC099023 Drosophi
32	225.4	5.9	180814	3 AC099012	AC099012 Drosophi
33	225.4	5.9	183592	3 AE003792	AE003792 Drosophi
34	225.4	5.9	256764	3 S80110	S80110 hormone-sen
35	186.6	4.9	225	4 AC073803	AC073803 Mus muscu
36	171	4.5	190860	2 AY051867	AY051867 Drosophi
37	149.8	3.9	2328	3 AC109038	AC109038 Rattus no
38	112.8	3.0	142394	2 D90903	D90903 Synchocyst
39	102	2.7	106209	1 PFATRAA	M17877 P.falciparu
40	99.8	2.6	2359	3 A00663	A00663 P.falciparu
41	96.6	2.5	3187	6 AC113928	AC113928 Rattus no
42	96.4	2.5	137721	2 ZMPROPR	Y17332 Zea mays MR
43	94.6	2.5	1471	8 ZMA130830	AJ130830 Zea mays
44	94.6	2.5	1792	8 AY061106	AY061106 Drosophi
45	87.4	2.3	4186	3	

ALIGNMENTS

RESULT 1

AX113509

LOCUS AX113509 3804 bp

DEFINITION Sequence 1 from Patent WO0126664.

ACCESSION AX113509

VERSION AX113509.1 GI:13939727

KEYWORDS

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 3804)

AUTHORS Mitchell,G.A. and Wang,S.P.

TITLE Hormone-sensitive lipase activity mediated male infertility

JOURNAL Patent: WO 0126664-A 1 19-Apr-2001;

Pred. No. is the number of results predicted by chance to have a

FEATURES	Hopital Sainte-Justine (CA)
source	Location/Qualifiers
	1. .3804
	/organism="Homo sapiens"
	/db.xref="taxon:9606"
	278. .3508
CDS	/note="unnamed protein product"
	/codon_start=1
	/protein_id="CAC37855.1"
	/db.xref="GI:13939728"
	/translation="MEPKSVSRSDMOPPEPHORITPLEPEKPTPIAPESKTLQGSNTQKPAHQNPILQOETPAQDAESQKERRAQKSAQOEFLAPQKPAQOQSPYIQKVLITQDEASQOQPGELGKESITQOEPALRQVGPQGPPEPAQOEAESTPAQAKPGARESPAPESTSOETPEQSDKOTTPVOGAKKQOSLFTGLTLOESTORSALEMKALSEWVADSESDVSSDSTPATMGAKQOQKGLGFKKSGYKXMSGSYSGTSPEKTSARNHHYQDTSARLIHNMDELPMOSTVTLAEDNIAFPSSQPEPTQRLSGVFQREDAQGLGELPALGRILGVAILPDLDETPPANGYRLVHTARCLALHLKRSYVANSRSTFPPRSHNLAEVLALITOLRALVYAOQLVTHNRPGLVFEGBGLTADPLREYVTLHKRCFYGRCLGFQFPALPPLQITISGLVSEGHYKRNETGLSYAASSLFTSGRAIDELRGAEEFRTIONLDVHFMKAENMITEMEYLSIANNASATVRSLSISLEPEAEMLADPPLVLTISPPLATGPGVPLVRLISYDLREGDSELSLIKSNNGORLELMPPOQAPRSRLIHFHGGGFAVOTSRSEPTIKSMAGOBAPITISIDYSLAPEAPPEALRECEFFAYCWAIKHCALIGSTGRICLAGSAGAGNCTTVALPAAAYGVRPDGIIMAYPATMLQAPASPERLSLMDPLPLVLSKVCVSAVYAKTEHNSDOKALGMELVRRDRLALDRFLRCLASVLSNLSFLSGRSQKMEPIAEPMRVSSEAAALADQOPLGTDLSLKNLRLDLRSLSNGSESSDPEMSLSLETLSPTSPVNLIPPDDAGEAEAKNELSPMDRLGVRAPAEFGFPRRRSOGATPMPLSYSPYKNPMSPLLADDSMKSLPPIHIVACALDPMLDSDSVMALARLRNLGQPTPLFVVDLPHGFLTLAALCRETROAEELCEVRIRLVLPFPAGAGPSGETGAAGVDDGGCGGNH"
BASE COUNT	828 a 1225 c 1091 g 660 t
ORIGIN	
Query Match	100.0%; Score 3804; DB 6; Length 3804;
Best Local Similarity	100.0%; Pred. No. 0;
Matches 3804; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	1 CTTCTGTGAAGAGTGTAGGACATACCCCTCTTCTTCTTAATCTCCACCAAG 60
DB	1 CTTCTGTGAAGAGTGTAGGACATACCCCTCTTCTTCTTAATCTCCACCAAG 60
QY	61 AAAGAGGACAGAGTTCATTAGTGGGGCCAGCTGTGATCGGCCAATCGCCAGCTG 120
DB	61 AAAGAGGACAGAGTTCATTAGTGGGGCCAGCTGTGATCGGCCAATCGCCAGCTG 120
QY	121 CCTTAAAAAGAGACCAAGTGTGTAGAGTGAATGAACCCCAAGAGAAATGCCATCA 180
DB	121 CCTTAAAAAGAGACCAAGTGTGTAGAGTGAATGAACCCCAAGAGAAATGCCATCA 180
QY	181 TGAGGAATCAATGAGAGTCTGTGAAGAGAGAGGCTGGGTGGAGCCCAAGAGATGA 240
DB	181 TGAGGAATCAATGAGAGTCTGTGAAGAGAGAGGCTGGGTGGAGCCCAAGAGATGA 240
QY	241 ACCTGGAAGTCAATATCTCTCCCTGTGAGGGAATTAACAATGAGAGCCAGTTCTAAGTCAGT 300
DB	241 ACCTGGAAGTCAATATCTCTCCCTGTGAGGGAATTAACAATGAGAGCCAGTTCTAAGTCAGT 300
QY	301 GTTCTAGGTCAAGTGTGCAACCTGGAACCAACAGAGGCTTAAACCCGCTAGAGCTGG 360
DB	301 GTTCTAGGTCAAGTGTGCAACCTGGAACCAACAGAGGCTTAAACCCGCTAGAGCTGG 360
QY	361 GCCAGAAAAAGACACCTATAGCCAGCAGATCGAAGACTGTGAGGAGATTCATTAACCA 420
DB	361 GCCAGAAAAAGACACCTATAGCCAGCAGATCGAAGACTGTGAGGAGATTCATTAACCA 420
QY	421 ACAGAAAGCTGCTTCAAAACCAAGACCCCTACCCAGCAGAGAGACCCCTGCACAAATGA 480
DB	421 ACAGAAAGCTGCTTCAAAACCAAGACCCCTACCCAGCAGAGAGACCCCTGCACAAATGA 480
QY	481 TGCTGAATCCAGAGAGACCTAGAGCCCAACAAATCTGCTTCAACAAAGAAATTTCT 540
DB	481 TGCTGAATCCAGAGAGACCTAGAGCCCAACAAATCTGCTTCAACAAAGAAATTTCT 540
QY	541 TGCCCCACAGAGAGCCGACACAGCAATCACTTACATCAAAAGGATGCTGCTACTCA 600

DB	541 TGCCCCACAGAGAGCCGACACAGCAATCACTTACATCAAAAGGATGCTGCTACTCA 600
QY	601 ACAGGAAGCTGCTCCAGAGGAGACCTGGCTAGAGAAAAATCTATATCTACACAGA 660
DB	601 ACAGGAAGCTGCTCCAGAGGAGACCTGGCTAGAGAAAAATCTATATCTACACAGA 660
QY	661 GCCAGATTGAGCAAAAGACATGTAGCCAGCCAGGAGGCTGGGCCAGAGAGCCACTCC 720
DB	661 GCCAGATTGAGCAAAAGACATGTAGCCAGCCAGGAGGCTGGGCCAGAGAGCCACTCC 720
QY	721 AGCTCAACAAAGAGCTGAATCAACACCTGGGCCAGGCTAAACCTGAGCCAAAAGGA 780
DB	721 AGCTCAACAAAGAGCTGAATCAACACCTGGGCCAGGCTAAACCTGAGCCAAAAGGA 780
QY	781 GCATCTGGCCCGAGCTGAATCTACATCCCAAGAGACACTGTAACAGTGCAGACCAAC 840
DB	781 GCATCTGGCCCGAGCTGAATCTACATCCCAAGAGACACTGTAACAGTGCAGACCAAC 840
QY	841 AACGCCAGTCCAGGAGCCAAATCCAGAGGAGATCTTGACAGAGCTGGATTTCTAAC 900
DB	841 AACGCCAGTCCAGGAGCCAAATCCAGAGGAGATCTTGACAGAGCTGGATTTCTAAC 900
QY	901 AAACCTTCAGAACTATCTCATACAGCGATCAAGCCCTAGAGTGAAGGCACTTTCTGAGTG 960
DB	901 AAACCTTCAGAACTATCTCATACAGCGATCAAGCCCTAGAGTGAAGGCACTTTCTGAGTG 960
QY	961 GGTGCAGATTCGTGATGCAGATCAAGATGTGGATCATCTTCAGACACAGATTCGCCAGC 1020
DB	961 GGTGCAGATTCGTGATGCAGATCAAGATGTGGATCATCTTCAGACACAGATTCGCCAGC 1020
QY	1021 CACGATGGGTGGAATGTGTGGCCCAAGGAGTGAAGCTTCAAGAGAAATCTGTTA 1080
DB	1021 CACGATGGGTGGAATGTGTGGCCCAAGGAGTGAAGCTTCAAGAGAAATCTGTTA 1080
QY	1081 TAAAGTGTGTAGAGTACAGTGTGGAGAGTGCACATGAGAAAAACAGTGTGCGAATCA 1140
DB	1081 TAAAGTGTGTAGAGTACAGTGTGGAGAGTGCACATGAGAAAAACAGTGTGCGAATCA 1140
QY	1141 CAGACACTACAGAGATACAGCTTCAAGGCTCATCCACAACATGAGACCTGGGCACAAATGAC 1200
DB	1141 CAGACACTACAGAGATACAGCTTCAAGGCTCATCCACAACATGAGACCTGGGCACAAATGAC 1200
QY	1201 ACAGTGGTGTGACTGTGGGGGAGGACAAATAGCTTCTTCTGAGAGCAAGGCTCTGG 1260
DB	1201 ACAGTGGTGTGACTGTGGGGGAGGACAAATAGCTTCTTCTGAGAGCAAGGCTCTGG 1260
QY	1261 GGAAGCGGCCAGCGGCTGTCAAGCGCTTTTGGCGGTGACAGGAGAGGCGCTGGGGCT 1320
DB	1261 GGAAGCGGCCAGCGGCTGTCAAGCGCTTTTGGCGGTGACAGGAGAGGCGCTGGGGCT 1320
QY	1321 GGAGCGGGCCCTGGGCGGCTGTGGGTGTGGCGACCTCTTGTGACCTGGACCCAGAGAC 1380
DB	1321 GGAGCGGGCCCTGGGCGGCTGTGGGTGTGGCGACCTCTTGTGACCTGGACCCAGAGAC 1380
QY	1381 ACCGGCCAAAGGAGTACGCGAGCTAGTGCACACAGCCGCTGTGCTGGCGAGACCTCT 1440
DB	1381 ACCGGCCAAAGGAGTACGCGAGCTAGTGCACACAGCCGCTGTGCTGGCGAGACCTCT 1440
QY	1441 GCACAAATCCCGTATGTGGCTTCAACCGCGCAGACATCTTCTTCGACACAGCCACAA 1500
DB	1441 GCACAAATCCCGTATGTGGCTTCAACCGCGCAGACATCTTCTTCGACACAGCCACAA 1500
QY	1501 CCTGGCGGAGCTGGAGGCTTACCTGGCTGCTCAACCCAGCTCCGCGCTGTGTTACTTA 1560
DB	1501 CCTGGCGGAGCTGGAGGCTTACCTGGCTGCTCAACCCAGCTCCGCGCTGTGTTACTTA 1560
QY	1561 CGCCAGCGCTGCTGTGTACCAATCGCGCGGGGTACTCTTCTTGAAGGCGAGAGGG 1620
DB	1561 CGCCAGCGCTGCTGTGTACCAATCGCGCGGGGTACTCTTCTTGAAGGCGAGAGGG 1620
QY	1621 GCTCACCGCGACTTCTCTCGGAGATGTCAAGCTGCATAAGGAGATCTTCTATGGCGG 1680

```
Db 1621 GCTACACGCCGACTCTCCGGGAGTATGTACGCTGCATAAAGGATGTTCTATGCGC 1680
QY 1681 CTGCGTGGGCTTCAGTTACGCTGCATCGCGCCATTCTCTGAGACCAATTCATTGG 1740
Db 1681 CTGCGTGGGCTTCAGTTACGCTGCATCGCGCCATTCTCTGAGACCAATTCATTGG 1740
QY 1741 GCTGTGTCCTTCGGGGAGCACTACAAACGCAACGAGACAGGCTCTAGTGTGGCGCAG 1800
Db 1741 GCTGTGTCCTTCGGGGAGCACTACAAACGCAACGAGACAGGCTCTAGTGTGGCGCAG 1800
QY 1801 CTCTCTCTTACACGCGCGCTTTTGCCATCGACCCGAGCTGCTGGGGCTGATTGA 1860
Db 1801 CTCTCTCTTACACGCGCGCTTTTGCCATCGACCCGAGCTGCTGGGGCTGATTGA 1860
QY 1861 GCGGATCACACAGAACCTGGAGCTGCATCTTGAAAAGCTTCTGSAACATCACCGGAT 1920
Db 1861 GCGGATCACACAGAACCTGGAGCTGCATCTTGAAAAGCTTCTGSAACATCACCGGAT 1920
QY 1921 GGAAGTGTATCTCTCTGGCCAAATGGCATGCGCCAGCTGAGGGTAAAGCCGCTGCT 1980
Db 1921 GGAAGTGTATCTCTCTGGCCAAATGGCATGCGCCAGCTGAGGGTAAAGCCGCTGCT 1980
QY 1981 CAGCTTCGACCCGAAACCTTTAGATGCCACTGACTGCGGACCCCAAGCTCACGGTAC 2040
Db 1981 CAGCTTCGACCCGAAACCTTTAGATGCCACTGACTGCGGACCCCAAGCTCACGGTAC 2040
QY 2041 CATCTACACCCCACTGGCCACACAGGCCCTGGGCGCGTCTGTCAGGCTCATCTCTA 2100
Db 2041 CATCTACACCCCACTGGCCACACAGGCCCTGGGCGCGTCTGTCAGGCTCATCTCTA 2100
QY 2101 TGAACCTGCTGAAGAGACAGACAGTGAAGAGCTCAGCAGCTGTAAAGTCAACGGCCA 2160
Db 2101 TGAACCTGCTGAAGAGACAGACAGTGAAGAGCTCAGCAGCTGTAAAGTCAACGGCCA 2160
QY 2161 AGGAGACCTTGAGACTGTGCGCGCGCCGCCAGAGGACCCCGCTCGCGTCTGATAGT 2220
Db 2161 AGGAGACCTTGAGACTGTGCGCGCGCCGCCAGAGGACCCCGCTCGCGTCTGATAGT 2220
QY 2221 GCACCTTCACGGGCGTGGCTTTGTGGCCAGACTCCAGATCCAGACGACCTTAACCTCAA 2280
Db 2221 GCACCTTCACGGGCGTGGCTTTGTGGCCAGACTCCAGATCCAGACGACCTTAACCTCAA 2280
QY 2281 GAGCTGGGCCACAGAGACTGGGCGCCGCCATCATCTCATGACTACCTCCCTGCGCTGA 2340
Db 2281 GAGCTGGGCCACAGAGACTGGGCGCCGCCATCATCTCATGACTACCTCCCTGCGCTGA 2340
QY 2341 GGGCCCTTTCGCCGTGGAGAGTCTTCTGCTACTGCTGGCCATCAAGCA 2400
Db 2341 GGGCCCTTTCGCCGTGGAGAGTCTTCTGCTACTGCTGGCCATCAAGCA 2400
QY 2401 CTGGCGCCCTCTTGGCTCAACAGGGGAGAAATCTGCTGGGGGAGACAGTCAAGGCG 2460
Db 2401 CTGGCGCCCTCTTGGCTCAACAGGGGAGAAATCTGCTGGGGGAGACAGTCAAGGCG 2460
QY 2461 GAACCTCTGCTTACCCGTGAGCTTCTGGGAGCAGCCTACGGGGTGGGGTGCAGATGG 2520
Db 2461 GAACCTCTGCTTACCCGTGAGCTTCTGGGAGCAGCCTACGGGGTGGGGTGCAGATGG 2520
QY 2521 CATCATGAGAGCTTACCCGCGCAATTCGTGACGCTGCGCGCTCTCCCTCCGCGCTGT 2580
Db 2521 CATCATGAGAGCTTACCCGCGCAATTCGTGACGCTGCGCGCTCTCCCTCCGCGCTGT 2580
QY 2581 GAGCCTCATGAGACCCCTTGTGCGCCCTCAAGTGTCTCCAAATGTCTACGCGCTATGCG 2640
Db 2581 GAGCCTCATGAGACCCCTTGTGCGCCCTCAAGTGTCTCCAAATGTCTACGCGCTATGCG 2640
QY 2641 TGTGCAAAAGAGAGAGACACTCAACTCAGACAGAAAGCCCTGCGCATGTAGTGGGCT 2700
Db 2641 TGTGCAAAAGAGAGAGACACTCAACTCAGACAGAAAGCCCTGCGCATGTAGTGGGCT 2700
QY 2701 GGTGGGGGGGAGACAGCCCTGCTCGGAGACTTCCGCTGGGTGCTCTCATGAGCT 2760
Db 2701 GGTGGGGGGGAGACAGCCCTGCTCGGAGACTTCCGCTGGGTGCTCTCATGAGCT 2760
QY 2761 CAACCTCTTCCGGAGTTAAAGTGGGCGCAAGTCCCAAGAAATGTGAGCCCATGACGA 2820
Db 2761 CAACCTCTTCCGGAGTTAAAGTGGGCGCAAGTCCCAAGAAATGTGAGCCCATGACGA 2820
QY 2821 GCCATGCGCGCGAGTGTGTGAAGCAGCACTGGCCAGCCCGAGGCGCCACTGGGAC 2880
Db 2821 GCCATGCGCGCGAGTGTGTGAAGCAGCACTGGCCAGCCCGAGGCGCCACTGGGAC 2880
QY 2881 GGATTCCTCAAGAACTGACCTGAGGACTTGAAGCTTGAAGGAAATCTCCGAGAGCTC 2940
Db 2881 GGATTCCTCAAGAACTGACCTGAGGACTTGAAGCTTGAAGGAAATCTCCGAGAGCTC 2940
QY 2941 GTGGAGACCCCGGAGTGTGCTGTCAGTGTGAGACACTTTGAGCCCTCCACACCCCTCGA 3000
Db 2941 GTGGAGACCCCGGAGTGTGCTGTCAGTGTGAGACACTTTGAGCCCTCCACACCCCTCGA 3000
QY 3001 TGTCAACTTCTTATTAACCACTGAGAGATGACAGGGGAGAGGCTGAGGCCAAATGAGCT 3060
Db 3001 TGTCAACTTCTTATTAACCACTGAGAGATGACAGGGGAGAGGCTGAGGCCAAATGAGCT 3060
QY 3061 GAGCCCATGAGACAGAGGCTTGGGCTCGTGCAGCTTCCCGAGGTTTCCACCCCG 3120
Db 3061 GAGCCCATGAGACAGAGGCTTGGGCTCGTGCAGCTTCCCGAGGTTTCCACCCCG 3120
QY 3121 ACGCTCCAGCGAGGTTGCCACAGAGATGCCCCCTACTCTCCTACCCATGATCAAGAACCC 3180
Db 3121 ACGCTCCAGCGAGGTTGCCACAGAGATGCCCCCTACTCTCCTACCCATGATCAAGAACCC 3180
QY 3181 CTTCATGTGCGCGCTGTGTCGACCCGACAGCATGCTCAAGAGCCCTGCCACCTGTGACAT 3240
Db 3181 CTTCATGTGCGCGCTGTGTCGACCCGACAGCATGCTCAAGAGCCCTGCCACCTGTGACAT 3240
QY 3241 CGTGGCGTGGCGCTGACCCCATGTGAGACACTCGGTCAATGCTCGCGCGGCACTGCG 3300
Db 3241 CGTGGCGTGGCGCTGACCCCATGTGAGACACTCGGTCAATGCTCGCGCGGCACTGCG 3300
QY 3301 CAACCTGGGCGACGCGGTGAGCGTGGCGGTGGTGAAGAACCTGCCGCGAGGCTTCTCTAG 3360
Db 3301 CAACCTGGGCGACGCGGTGAGCGTGGCGGTGGTGAAGAACCTGCCGCGAGGCTTCTCTAG 3360
QY 3361 CCTAGCGCGCTGTGTCGCGAGAGCGCGCAGGCGCAGAGCTGTGCTGAGAGCGCATCG 3420
Db 3361 CCTAGCGCGCTGTGTCGCGAGAGCGCGCAGGCGCAGAGCTGTGCTGAGAGCGCATCG 3420
QY 3421 CCTGTCTCTACTCTCCCGCGAGACCGGCGCAGCGGGAGAGAGCGGGGCTGCGGGGT 3480
Db 3421 CCTGTCTCTACTCTCCCGCGAGACCGGCGCAGCGGGAGAGAGCGGGGCTGCGGGGT 3480
QY 3481 AGAGGGGGGCTGGGGGGGAGCACTAAAGCTGTGTTCCATCTGCGCGCGGCTCGG 3540
Db 3481 AGAGGGGGGCTGGGGGGGAGCACTAAAGCTGTGTTCCATCTGCGCGCGGCTCGG 3540
QY 3541 TCATGAATGCTTCCGGGCGCGGCGGAGAGGAGCGGGCTGTGTTAAATGTCGGG 3600
Db 3541 TCATGAATGCTTCCGGGCGCGGCGGAGAGGAGCGGGCTGTGTTAAATGTCGGG 3600
QY 3601 GTGGCAAGGGGGCGGGGCGCGGAGAACTGAGACCTTGGCCACGCGGGAGGGAGCGC 3660
Db 3601 GTGGCAAGGGGGCGGGGCGCGGAGAACTGAGACCTTGGCCACGCGGGAGGGAGCGC 3660
QY 3661 GCACACACACCGGTACCGAGAACAGGCTGAGCTGACGCGCACCGGTGCTTTGCTGCTG 3720
Db 3661 GCACACACACCGGTACCGAGAACAGGCTGAGCTGAGCTGACGCGCACCGGTGCTTTGCTGCTG 3720
QY 3721 CTGCTGGGCGACCGCGCAGGAGCAGGGAGCTGCGCTCTCCCTTGAAGTTCGGTTGTTT 3780
Db 3721 CTGCTGGGCGACCGCGCAGGAGCAGGGAGCTGCGCTCTCCCTTGAAGTTCGGTTGTTT 3780
QY 3781 GTTGAATAAATATTTAATTA 3804
Db 3781 GTTGAATAAATATTTAATTA 3804
```


|||||
Db 1201 AAGTGGCTGTGCTATGCGGAGACAAATAGCTTCTTCGAGCCAGGGCTCTCG 1260
OY 1261 GGAAGAGGCGCCAGCGGCTGTGAGCGCTTTTGGCGGTGTACGGGAGAGCGCGGGGCT 1320
Db 1261 GGAAGAGGCGCCAGCGGCTGTGAGCGCTTTTGGCGGTGTACGGGAGAGCGCGGGGCT 1320
OY 1321 GGAAGAGGCGCGCTGTGAGCGCTGTGAGCGCTTTTGGCGGTGTACGGGAGAGCGCGGGGCT 1380
Db 1321 GGAAGAGGCGCGCTGTGAGCGCTGTGAGCGCTTTTGGCGGTGTACGGGAGAGCGCGGGGCT 1380
OY 1381 ACCGGGCAACGGGTACCGGAGCTAGTGACACAGCCGCTGCTGCTGGCGACCTTCCT 1440
Db 1381 ACCGGGCAACGGGTACCGGAGCTAGTGACACAGCCGCTGCTGCTGGCGACCTTCCT 1440
OY 1441 GCAAAATCCCGGTATGTGGCTTCAACCGCGCGAGCATTTCTCCGACACAGCCACAA 1500
Db 1441 GCAAAATCCCGGTATGTGGCTTCAACCGCGCGAGCATTTCTCCGACACAGCCACAA 1500
OY 1501 CCGGCGGAGCTGGAGGCGCTACCTGGCTGCTACCCAGCTCGCGCTGCTGCTACTA 1560
Db 1501 CCGGCGGAGCTGGAGGCGCTACCTGGCTGCTACCCAGCTCGCGCTGCTGCTACTA 1560
OY 1561 CGCGGAGCGCTGTGAGCTTACCAATCGCGGGGGTACTCTTTGAGGGCGACGAGGG 1620
Db 1561 CGCGGAGCGCTGTGAGCTTACCAATCGCGGGGGTACTCTTTGAGGGCGACGAGGG 1620
OY 1621 GCTCACCGCGGACTTCTCTCGGGAGTATGTACAGCTGCATAAAGGATGCTTATGGCCG 1680
Db 1621 GCTCACCGCGGACTTCTCTCGGGAGTATGTACAGCTGCATAAAGGATGCTTATGGCCG 1680
OY 1681 CTGCGCTGGGCTTCAGTTTCAGCGCTGCATCGGGCATTCCTGGAGACATCTCATTTGG 1740
Db 1681 CTGCGCTGGGCTTCAGTTTCAGCGCTGCATCGGGCATTCCTGGAGACATCTCATTTGG 1740
OY 1741 GCTGTGTCTTTCGGGAGCACTACAAACGCAAGAGACAGCGCTCAGTGTGGCGCGCAG 1800
Db 1741 GCTGTGTCTTTCGGGAGCACTACAAACGCAAGAGACAGCGCTCAGTGTGGCGCGCAG 1800
OY 1801 CTCTCTTTCACACGAGGCGCGCTTTCGCATCGACCCGAGACTGCTGTGGGCTAGTTTGA 1860
Db 1801 CTCTCTTTCACACGAGGCGCGCTTTCGCATCGACCCGAGACTGCTGTGGGCTAGTTTGA 1860
OY 1861 GCGGATACACAGAAACCTGGAGCTGCATCTTGAAAGCTTCTGGAACATCACCGAGAT 1920
Db 1861 GCGGATACACAGAAACCTGGAGCTGCATCTTGAAAGCTTCTGGAACATCACCGAGAT 1920
OY 1921 GGAAGTGTATGCTCTGTGGCAACATGCGATCGGCCACCGTGAAGGCTAAAGCCGCTGT 1980
Db 1921 GGAAGTGTATGCTCTGTGGCAACATGCGATCGGCCACCGTGAAGGCTAAAGCCGCTGT 1980
OY 1981 CAGCGTGGCACCCGAGAGCTTTAGATGCGCATGACCTGCGAGCCCGACGCTCACGGTAC 2040
Db 1981 CAGCGTGGCACCCGAGAGCTTTAGATGCGCATGACCTGCGAGCCCGACGCTCACGGTAC 2040
OY 2041 CATCTACCCCGCTGAGGCGCGACAGAGCCCTTGAGAGCTGACGACCGCCACGCTCACGGTAC 2100
Db 2041 CATCTACCCCGCTGAGGCGCGACAGAGCCCTTGAGAGCTGACGACCGCCACGCTCACGGTAC 2100
OY 2101 TGAAGTGTATGAG 2160
Db 2101 TGAAGTGTATGAG 2160
OY 2161 ACGGAGAGCTGGAGCTGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2220
Db 2161 ACGGAGAGCTGGAGCTGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2220
OY 2221 GCACTTCCACGCGGCGGTGTGGCGCGCGAGCTCCAGATCCAGAGAGCGCTTACCTCAA 2280
Db 2221 GCACTTCCACGCGGCGGTGTGGCGCGCGAGCTCCAGATCCAGAGAGCGCTTACCTCAA 2280
OY 2281 GAGCTGGGCGCGAGAGCTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2340
Db 2281 GAGCTGGGCGCGAGAGCTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2340

Db 2281 GAGCTGGGCGCGAGAGCTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2340
OY 2341 GCGCCCGCTTCCCGGTGGCGTGGAGAGTGTCTTCTGCTACTGCTGGCGCATCAAGCA 2400
Db 2341 GCGCCCGCTTCCCGGTGGCGTGGAGAGTGTCTTCTGCTACTGCTGGCGCATCAAGCA 2400
OY 2401 CTGGCGCGCTTCCCGGTGGCGTGGAGAGTGTCTTCTGCTACTGCTGGCGCATCAAGCA 2460
Db 2401 CTGGCGCGCTTCCCGGTGGCGTGGAGAGTGTCTTCTGCTACTGCTGGCGCATCAAGCA 2460
OY 2461 GAAGCTGTGCTTACCGGTGTCTTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2520
Db 2461 GAAGCTGTGCTTACCGGTGTCTTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2520
OY 2521 CATATGGAG 2580
Db 2521 CATATGGAG 2580
OY 2581 GAGCGTATGAG 2640
Db 2581 GAGCGTATGAG 2640
OY 2641 TGTGCAAAAG 2700
Db 2641 TGTGCAAAAG 2700
OY 2701 GGTGCGGCGGAGACAGAGCGCTTCTCGAGAGCTTCCGCTGGGTGCTCTCATGGCT 2760
Db 2701 GGTGCGGCGGAGACAGAGCGCTTCTCGAGAGCTTCCGCTGGGTGCTCTCATGGCT 2760
OY 2761 CAATCTCTCTGAGATTAAATGGGCGCAAGTCCAGAGAGATGTGAGAGCCATAGCAGA 2820
Db 2761 CAATCTCTCTGAGATTAAATGGGCGCAAGTCCAGAGAGATGTGAGAGCCATAGCAGA 2820
OY 2821 GCCATGCGCGCGAGTGTGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2880
Db 2821 GCCATGCGCGCGAGTGTGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2880
OY 2881 GGATTCCTCTAAGAAACCTGAG 2940
Db 2881 GGATTCCTCTAAGAAACCTGAG 2940
OY 2941 GTGCGAGACCCCGAGAGTGTGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3000
Db 2941 GTGCGAGACCCCGAGAGTGTGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3000
OY 3001 TGTCAACTTCTTATTAACCACTGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3060
Db 3001 TGTCAACTTCTTATTAACCACTGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3060
OY 3061 GAGCGCGATGAG 3120
Db 3061 GAGCGCGATGAG 3120
OY 3121 ACGCTTCCAGCGAG 3180
Db 3121 ACGCTTCCAGCGAG 3180
OY 3181 CTTCATGTGCGCGCTGTGCGAG 3240
Db 3181 CTTCATGTGCGCGCTGTGCGAG 3240
OY 3241 CGTGGCGTGGCGGTGGAG 3300
Db 3241 CGTGGCGTGGCGGTGGAG 3300
OY 3301 CAACCTGGGCGAGCGGAG 3360
Db 3301 CAACCTGGGCGAGCGGAG 3360
OY 3361 CTAAGCGGCGCTGTGCGCGAG 3420
Db 3361 CTAAGCGGCGCTGTGCGCGAG 3420

QY 3421 CCGTCCCTCACTCTCCCGCGAGCGGCGAGAGCGGGGCTGCGGGGT 3480
|||
Db 3421 CCTGCTCTCACTCTCCCGCGAGCGGCGAGAGCGGGGCTGCGGGGT 3480
QY 3481 AGACGGGGGCTGCGGGGCGACACTAAAGCTTGTGTCCATCTGCGCGCTCGG 3540
|||
Db 3481 AGACGGGGGCTGCGGGGCGACACTAAAGCTTGTGTCCATCTGCGCGCTCGG 3540
QY 3541 TCATCAATGCTTCCTCCGGGCGGGAAGGAGCGGGCTGCTTACTTAAGTCGGGG 3600
|||
Db 3541 TCATCAATGCTTCCTCCGGGCGGGAAGGAGCGGGCTGCTTACTTAAGTCGGGG 3600
QY 3601 GTGGCAAGGGGCGGGGCGGGAAGCTGAGACCTTCGCCACGAGGAGGAGACCC 3660
|||
Db 3601 GTGGCAAGGGGCGGGGCGGGAAGCTGAGACCTTCGCCACGAGGAGGAGACCC 3660
QY 3661 GCACACACACCGGCTACCGAGAGGCTGAGCCTGACAGCGCAGCTGCTTGTGCTG 3720
|||
Db 3661 GCACACACACCGGCTACCGAGAGGCTGAGCCTGACAGCGCAGCTGCTTGTGCTG 3720
QY 3721 CTGCTGGGCGGACCGCGGAGGAGGAGCTGCGCTTCCTTGCAGTGGTGTGTT 3780
|||
Db 3721 CTGCTGGGCGGACCGCGGAGGAGGAGCTGCGCTTCCTTGCAGTGGTGTGTT 3780
QY 3781 GTTGTAAATAAAGTATTTAATTA 3804
|||
Db 3781 GTTGTAAATAAAGTATTTAATTA 3804

RESULT 3
HUMHSLA
LOCUS HUMHSLA 3255 bp DNA linear PRI 13-JUL-1995
DEFINITION Human hormone-sensitive lipase (LIPe) gene, complete cds.
ACCESSION L11706
VERSION L11706.1 GI:896474
KEYWORDS hormone-sensitive lipase.
SOURCE Homo sapiens male adult DNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS 1 (bases 1 to 3255)
Langin,D., Laurell,H., Holst,L.S., Belfrage,P. and Holm,C.
Gene organization and primary structure of human hormone-sensitive
lipase: possible significance of a sequence homology with a lipase
of Moraxella TA144, an antarctic bacterium
Proc. Natl. Acad. Sci. U.S.A. 90 (11), 4897-4901 (1993)
JOURNAL
MEDLINE 9328163
PUBMED 8506334
REFERENCE 2 (bases 1 to 3255)
AUTHORS Holm,C.
TITLE Direct Submission
JOURNAL Submitted (02-MAR-1993) Cecilia Holm, Cell and Molecular Biology,
University of Lund, PO. Box 94, Lund, S-221 00, Sweden
On Jul 13, 1995 this sequence version replaced gi:895926.
COMMENT
FEATURES
source
1. 3255
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="19q13.1-q13.2"
/sex="male"
/cell_type="lymphocyte"
/dev_stage="adult"
1. 1147
/gene="LIPe"
/product="hormone-sensitive lipase"
/note="G00-120-154"
/number=1
1. 631
/gene="LIPe"
/note="G00-120-154"
join(632..1147,1148..1238,1239..1384,1385..1570,
1571..1865,1866..2093,2094..2270,2271..2695,2696..2959)

CDS
/gene="LIPe"
join(632..1147,1148..1238,1239..1384,1385..1570,
1571..1865,1866..2093,2094..2270,2271..2695,2696..2959)
/gene="LIPe"
/codon_start=1
/product="hormone-sensitive lipase"
/protein_id="AAA69810.1"
/db_xref="GI:896476"
/db_xref="GDB:G00-120-154"
/translation="MDLRPMQSLVTLADNRIAFRSQGPGETAQRSLGVAFVREQA
LGLPRLRLGLVAHLEFDLDETPANGRLSVLHTRRCCIALHLHRSRYASRSRIFP
RTSHNLAELEVALAQLTQRLALVYAQRLVYNRGVLEFEGDESLADFLPEVTLH
KCFEYGRIGQEPFPAIRPFLQTLISGLVSEFHEHYKRNDETGLSVASSLFTSGRAID
PLRCAEFERITONLDVHFMAFVNITMEVLSLANNASATRVSRLLSLPEAFEM
PILADPTLVITISPLAHTGPGPVIVRLISYDLRGDSEELSLIKSGORLEMP
RQQAQRSRSLVHFRGGFVAQTSRSHPRYKSWAOELGAIIISIDYSLAEAPFR
ALEECFAVCWAKIKCALIGSTGERICLAGDSAGNLCETVALRAAATGVPRDSIMA
AYPATMLOPASPSRLSLMDPLPLSVLSKCVSAVYAAKTEDHSNDOAKLGMGLV
RDRFTALLRDFRLGASWLNSEFLSLGSKSQMSDEPIAPMRYSSEALAPQPGIG
TDSLKNLRLRDLISLGNSETSDPTSEMLSAETLSPSPSDVNFLLPEDAGEAEAK
NELSPDRGIGVRAAFPEGPFHRRSSGATOMPISSPYVKNPFMSPLPADPSMKSIL
PVHTVACALDPMILDDSYMLARLRLNLGOPYTLRYEDLPHQFLTLALCRTRQAE
LCVERIRLVITPPRAGPSGETGAGVGGCGGRH"
1148..1238
/gene="LIPe"
/product="hormone-sensitive lipase"
/note="G00-120-154"
/number=2
1239..1384
/gene="LIPe"
/product="hormone-sensitive lipase"
/note="G00-120-154"
/number=3
1385..1570
/gene="LIPe"
/product="hormone-sensitive lipase"
/note="G00-120-154"
/number=4
join(1571..1633,1637..1865)
/gene="LIPe"
/product="hormone-sensitive lipase"
/note="G00-120-154"
/number=5
1866..2093
/gene="LIPe"
/product="hormone-sensitive lipase"
/note="G00-120-154"
/number=6
2094..2270
/gene="LIPe"
/product="hormone-sensitive lipase"
/note="G00-120-154"
/number=7
2271..2695
/gene="LIPe"
/product="hormone-sensitive lipase"
/note="G00-120-154"
/number=8
join(2696..2778,2781..2837,2839..3255)
/gene="LIPe"
/product="hormone-sensitive lipase"
/note="G00-120-154"
/number=9
2960..3255
/gene="LIPe"
/note="G00-120-154"
3238..3243
/gene="LIPe"
/note="G00-120-154"
3243
/gene="LIPe"
/note="G00-120-154"
BASE COUNT 610 a 1057 c 1015 g 573 t
ORIGIN
Query Match 69.7%; Score 2650.4; DB 9; Length 3255;

Best Local Similarity 99.8%; Pred. No. 0;
Matches 2654; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1145 CACACACAGATAGAGCTCAAGCTCATCCACACATGAGCTCGGACATGACACAG 1204
| | | | |
Db 596 CCCACCTGGCCACAGGCTCAAGGCTCATCCAAATGAGCTCGGACATGACACAG 655
QY 1205 TCGCTGATGACTCGGGGAGAGACACATAGCTCTTCTGAGCCAGAGGCTCGGGAA 1264
| | | | |
Db 656 TCGCTGATGACTCGGGGAGAGACACATAGCTCTTCTGAGCCAGAGGCTCGGGAA 715
QY 1265 ACGGCCCCAGGCTGTGAGGCGTTTGGCGGTGTAGCGGAGCAGGCGCTGGAG 1324
| | | | |
Db 716 ACGGCCCCAGGCGTGTGAGGCGTTTGGCGGTGTAGCGGAGCAGGCGCTGGAG 775
QY 1325 CCGGCGCTGGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1384
| | | | |
Db 776 CCGGCGCTGGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 835
QY 1385 GCCAACGGGTACCGCAGCTAGTGCACACAGCCGCTGCTGCTGCTGCTGCTGCTG 1444
| | | | |
Db 836 GCCAACGGGTACCGCAGCTAGTGCACACAGCCGCTGCTGCTGCTGCTGCTGCTG 895
QY 1445 AAATCCGCTATGTGCGCTCCAAACCGCGCAGCATCTTCTTCCGACACGACAACTG 1504
| | | | |
Db 896 AAATCCGCTATGTGCGCTCCAAACCGCGCAGCATCTTCTTCCGACACGACAACTG 955
QY 1505 GCGGAGTGTGAGGCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1564
| | | | |
Db 956 GCGGAGTGTGAGGCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1015
QY 1565 CAGCGCTGTGCTGCTTACCAATCGCGCGGGGTACTCTTCTTGTAGGGGAGAGGGGCTC 1624
| | | | |
Db 1016 CAGCGCTGTGCTGCTTACCAATCGCGCGGGGTACTCTTCTTGTAGGGGAGAGGGGCTC 1075
QY 1625 ACCGCGCATCTCTCGGGAGATGTACAGCGCTGATTAAGGGATGCTTATGCGCGCTGC 1684
| | | | |
Db 1076 ACCGCGCATCTCTCGGGAGATGTACAGCGCTGATTAAGGGATGCTTATGCGCGCTGC 1135
QY 1685 CTGGGCTTCCAGTTCAAGCTGCGCATCGCGCATCTTCTGACAGACATCTCCATTGGGCTG 1744
| | | | |
Db 1136 CTGGGCTTCCAGTTCAAGCTGCGCATCGCGCATCTTCTGACAGACATCTCCATTGGGCTG 1195
QY 1745 GTGTCTTCTGGGAGAGCACTACAAAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1804
| | | | |
Db 1196 GTGTCTTCTGGGAGAGCACTACAAAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1255
QY 1805 CTCTTACACAGCGCGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1864
| | | | |
Db 1256 CTCTTACACAGCGCGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1315
QY 1865 ATACACAG 1924
| | | | |
Db 1316 ATACACAG 1375
QY 1925 GTGTATCTGTCTGTGCGCAATGATGCGCGCAGCTGAGGAGTAAAGCGCTGCTGAGC 1984
| | | | |
Db 1376 GTGTATCTGTCTGTGCGCAATGATGCGCGCAGCTGAGGAGTAAAGCGCTGCTGAGC 1435
QY 1985 CTGCGACCGAGAGCTTGTGAGATGCACTGACCTGCGACCCAGCTGACGGTCAACATC 2044
| | | | |
Db 1436 CTGCGACCGAGAGCTTGTGAGATGCACTGACCTGCGACCCAGCTGACGGTCAACATC 1495
QY 2045 TCACCCCGACCTGGCGCCACAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2104
| | | | |
Db 1496 TCACCCCGACCTGGCGCCACAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1555
QY 2105 CTGCGTGAAG 2164
| | | | |
Db 1556 CTGCGTGAAG 1615
QY 2165 AGCTGTGAGCTGTGCG 2224
| | | | |

Db 1616 AGCTGTGAGCTGTGCG 1675
QY 2225 TTCCACGGCGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2284
| | | | |
Db 1676 TTCCACGGCGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1735
QY 2285 TGGCGCCAGAGAGCTGGGCG 2344
| | | | |
Db 1736 TGGCGCCAGAGAGCTGGGCG 1795
QY 2345 CCGTTCCCGCGTGGGAGAGTGTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2404
| | | | |
Db 1796 CCGTTCCCGCGTGGGAGAGTGTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1855
QY 2405 GCGCTCTTGGCTCAACAGGGAGAACGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2464
| | | | |
Db 1856 GCGCTCTTGGCTCAACAGGGAGAACGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1915
QY 2465 CTCTGCTTACCGTGGCTCTTCTGCGGACAGAGCTACGGGGTTCGGGTGCGAGATGGCATC 2524
| | | | |
Db 1916 CTCTGCTTACCGTGGCTCTTCTGCGGACAGAGCTACGGGGTTCGGGTGCGAGATGGCATC 1975
QY 2525 ATGGAGCCTACCGGCGCAATGCTGAGCTGCGGCTCTCCCTCCGCGCTGCTGAGC 2584
| | | | |
Db 1976 ATGGAGCCTACCGGCGCAATGCTGAGCTGCGGCTCTCCCTCCGCGCTGCTGAGC 2035
QY 2585 CTGATGAGACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2644
| | | | |
Db 2036 CTGATGAGACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2095
QY 2645 GCAAG 2704
| | | | |
Db 2096 GCAAG 2155
QY 2705 CCGCGGAGACAGCGCTGCTCTCCGAGACTTCCGCTGCGGTGCTCTCATGCTCAAC 2764
| | | | |
Db 2156 CCGCGGAGACAGCGCTGCTCTCCGAGACTTCCGAGACTTCCGCTGCGGTGCTCTCATGCTCAAC 2215
QY 2765 TCCCTTCCGAGAGTAAAGTGGGCGCAAGTCCAGAGAGTGTGGAGAGCCATAGCAGAGCG 2824
| | | | |
Db 2216 TCCCTTCCGAGAGTAAAGTGGGCGCAAGTCCAGAGAGTGTGGAGAGCCATAGCAGAGCG 2275
QY 2825 ATGCGCCGACAGTGTGTGAG 2884
| | | | |
Db 2276 ATGCGCCGACAGTGTGTGAG 2335
QY 2885 TCCCTCAAGAACTGACCTTGAAGGACTTGAAGCTTGAAGGAGAACTCCGAGAGCTGCTG 2944
| | | | |
Db 2336 TCCCTCAAGAACTGACCTTGAAGGACTTGAAGCTTGAAGGAGAACTCCGAGAGAGCTGCTG 2395
QY 2945 GACACCCCGAGATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3004
| | | | |
Db 2396 GACACCCCGAGATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2455
QY 3005 AACTTCTTATTAACACCTGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3064
| | | | |
Db 2456 AACTTCTTATTAACACCTGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2515
QY 3065 CCCATGACAGAGAGCTGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3124
| | | | |
Db 2516 CCCATGACAGAGAGCTGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2575
QY 3125 TCCAGCCAGGGGTGCGACAGATGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3184
| | | | |
Db 2576 TCCAGCCAGGGGTGCGACAGATGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2635
QY 3185 ATGTGCGCGCTGTGCGACCCGACAGATGCTCAAGAGCTTCCACTGTGACATCTGTG 3244
| | | | |
Db 2636 ATGTGCGCGCTGTGCGACCCGACAGATGCTCAAGAGCTTCCACTGTGACATCTGTG 2695
QY 3245 GCGTGGCGGCTGTGAGCCCATGTGAGAGACTGTGCTATGCTGTGCGGCGGAGCTGTGCAAC 3304
| | | | |
Db 2696 GCGTGGCGGCTGTGAGCCCATGTGAGAGACTGTGCTATGCTGTGCGGCGGAGCTGTGCAAC 2755

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 3, 2003, 02:59:45 ; Search time 499 Seconds
(without alignments)
17167.538 Million cell updates/sec

Title: US-09-915-814-3

Perfect score: 3804
Sequence: 1 cctcttgtaagagagtgcta.....taataaagatttaatta 3804

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08
Maximum Match 1008
Listing first 45 summaries

Database : N_Geneseq.101002.*
1: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3804	100.0	3804	22	AAD05012 Human mRNA encodin
2	3227.8	84.9	3231	22	ABK87218 Human lipase, horm
3	1152.8	30.3	30709	24	ABK87217 Human lipase, horm
4	348.4	9.2	709	24	ABQ25066 Oligonucleotide fo
5	348.4	9.2	709	24	ABQ25067 Oligonucleotide fo
6	307	8.1	295	22	AA134322 Probe #12108 used
7	295	7.8	243	22	AA156402 Probe #25088 used
8	285.6	7.5	709	24	ABQ25068 Oligonucleotide fo
9	285.6	7.5	709	24	ABQ25069 Oligonucleotide fo

10	254	6.7	315	21	AAC08574 Human secreted
11	225.4	5.9	2752	23	ABL17205 Drosophila melanog
12	225.4	5.9	2781	23	ABL17207 Drosophila melanog
13	225.4	5.9	5133	23	ABL17206 Drosophila melanog
14	225.4	5.9	5307	23	ABL17204 Drosophila melanog
15	151.6	4.0	3720	24	ABK87221 Human lipase, horm
16	147.6	3.9	213	24	ABN18019 Human ORFX polynuc
17	110.4	2.9	217	21	AAA89608 Mouse hormone-sens
18	96.6	2.5	2790	7	AA60473 Sequence encoding
19	79.8	2.1	954	24	ABR91410 Oll-degrading bact
20	76.6	2.0	2885	24	ABL08459 Drosophila melanog
21	76.6	2.0	3848	23	ABL08136 Drosophila melanog
22	76.6	2.0	5004	23	ABL08458 pDIP/PyCSP.1 plasm
23	75	2.0	5552	24	AAA58309 Sugarcane proline
24	74.6	2.0	2548	24	AAD23837 Sugarcane proline
25	74.2	2.0	1172	24	AAD23835 Human secreted pro
26	70	1.8	230	21	ACC12553 Mycobacterium tube
27	69.4	1.8	4411529	22	AA199682 Insert from clone
28	65.4	1.7	1548	12	AAQ13868 Sugarcane proline
29	65	1.7	590	24	AAD23836 PRP 378 gene and r
30	64.8	1.7	3526	12	AAQ13864 DNA encoding whale
31	61	1.6	914	18	AAT79336 DNA encoding Arch
32	60.8	1.6	936	18	AAT79329 Human spliced tran
33	60	1.6	60	24	ABN32626 Human adenosine Al
34	58.4	1.5	114955	20	AAK53491 Human colon cancer
35	57.8	1.5	1127	21	AAA02477 Human encoding Met
36	55.8	1.5	905	18	AAT79333 Cell wall protein
37	55	1.4	5406	21	AAZ93533 Staph. epidermidis
38	55	1.4	5406	21	AAZ51201 Drosophila melanog
39	55	1.4	6353	23	ABL07173 Drosophila melanog
40	55	1.4	8429	23	ABL07172 Staphylococcus epi
41	54.6	1.4	3600	19	AAV04279 Human adenosine Al
42	53.8	1.4	114955	20	AAK53491 Drosophila melanog
43	52.6	1.4	3840	23	ABL18139 Drosophila melanog
44	52.6	1.4	7720	23	ABL18138 Human CDNA differe
45	52.6	1.4	8841	24	ABK84005

ALIGNMENTS

RESULT 1	
AAD05012	standard; mRNA; 3804 BP.
XX	
AC	AAD05012:
XX	
DT	17-JUL-2001 (first entry)
XX	
DE	Human mRNA encoding testicular isoform of hormone sensitive lipase.
XX	
KW	Human: hormone sensitive lipase; contraceptive; antisense-therapy;
KW	Inhibitor; testicular isoform; male fertility; ss.
XX	
OS	Homo sapiens.
XX	
FH	Key
FT	CDS
FT	Location/Qualifiers
FT	278..3508
FT	/*tag= a
FT	/product= "Human hormone sensitive lipase"
FT	/EC_number= "3.1.1.3"
XX	
PD	WO200126664-A2.
XX	
PD	19-APR-2001.
XX	
PF	13-OCT-2000; 2000WO-CA01228.
XX	
PR	14-OCT-1999; 99CA-2286451.
XX	
PA	(HOPI-) HOPITAL SAINTE-JUSTINE.
XX	
PI	Mitchell GA, Wang SP;

1741 GUGUGUGUCUUGGGAGACUACAAGCAACGAGACGGCTUCAGUGUGCCGCGAC 1800
1801 CTCCTCTTACACAGCGCGCTTTGCCATCAGCCCCGAGCTCGTGGGCTAGATTGA 1860
1801 CUCUCUCUACACAGCGCGCTTUCGCAVCGACCCGAGCTUCGUGGGCTAGGUGA 1860
1861 GGGATCACACAGAACCTGGACGTGACATTCTGAAAAGCTTGTGGAATCATCAGGAT 1920
1861 GCGGATACACAGAACCTUGACGUCGACUUCUGAAAAGCTTUCGACACUCCGAGAU 1920
1921 GGAAGTCTATGCTCTGTGGCCAATGCGATCGCCACCGTAGAGGTAAAGCCGCTCT 1980
1921 GGAAGUGUCUUGUCUGCCAAAVGGCAUGGCAACCGAGGUGAAGGUGCCGCTGCU 1980
1981 CAGCTCCACCCGAGAGCTTTGAGATGCATGACTGCCGACCCACGCTCAGGTCAC 2040
1981 CAGCTCCACCCGAGAGCTTTGAGATGCATGACTGCCGACCCACGCTCAGGTCAC 2040
2041 CATCTCACCCGCTGAGCCACAGAGGCGCTGGGCGCTGCTGTGACGGCTCATCTCTA 2100
2041 CAUUCACCCGCTGAGCCACAGAGGCGCTGGGCGCTGCTGTGACGGCTCATCTCTA 2100
2101 TGAAGTCTGAGAGAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2160
2101 UGACGUGGUGAAG 2160
2161 AGGAGCTGAGAGCTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 2220
2161 AGGAGCTGAGAGCTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 2220
2221 GCACTTCCAGAGGCGGTGGCTTGTGGCCGAGCTCCAGATCCAGAGGCTTACCTCA 2280
2221 GCACTTCCAGAGGCGGTGGCTTGTGGCCGAGCTCCAGATCCAGAGGCTTACCTCA 2280
2281 GAGCTGGGCGCAGAGAGCTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 2340
2281 GAGCTGGGCGCAGAGAGCTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 2340
2341 GGGCT 2400
2341 GGGCT 2400
2401 GCT 2460
2401 GCT 2460
2461 GAACTCTGCTTACAGCGGTGCTTGGGCGAGAGCTTACGGGCTGGGCTGGCAGATG 2520
2461 GAACTCTGCTTACAGCGGTGCTTGGGCGAGAGCTTACGGGCTGGGCTGGCAGATG 2520
2521 CATCATGAGAGCTTACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 2580
2521 CATCATGAGAGCTTACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 2580
2521 CATCATGAGAGCTTACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 2580
2521 CATCATGAGAGCTTACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 2580
2581 GAGCTCATGAGAGCTTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 2640
2581 GAGCTCATGAGAGCTTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 2640
2581 GAGCTCATGAGAGCTTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 2640
2581 GAGCTCATGAGAGCTTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 2640
2641 TGGTGCAG 2700
2641 TGGTGCAG 2700
2641 TGGTGCAG 2700
2641 TGGTGCAG 2700
2701 GGTGGCGCGCGAGACAGCGCTGCTCGAGAGCTTCCGCTGGGCTCGCTCATGCT 2760
2701 GGTGGCGCGCGAGACAGCGCTGCTCGAGAGCTTCCGCTGGGCTCGCTCATGCT 2760
2701 GGTGGCGCGCGAGACAGCGCTGCTCGAGAGCTTCCGCTGGGCTCGCTCATGCT 2760
2701 GGTGGCGCGCGAGACAGCGCTGCTCGAGAGCTTCCGCTGGGCTCGCTCATGCT 2760
2761 CAACCTCTCTGAGATTAAAGTGGGCGCAAGTCCAGAAAGATGCGGAGCCATAGCAG 2820
2761 CAACCTCTCTGAGATTAAAGTGGGCGCAAGTCCAGAAAGATGCGGAGCCATAGCAG 2820
2761 CAACCTCTCTGAGATTAAAGTGGGCGCAAGTCCAGAAAGATGCGGAGCCATAGCAG 2820
2821 GCGGATGCGCGCGAGTGTCTGAAGCAGCAGTGGCGCGCGCGCGCGCGCGCGCT 2880
2821 GCGGATGCGCGCGAGTGTCTGAAGCAGCAGTGGCGCGCGCGCGCGCGCGCGCT 2880

2821 GCGGATGCGCGCGAGUGUGUCGAGAGCAGCAGUGGCCAGCGCGCGCGCGCGCGCT 2880
2881 GATTTCCCTAAGAACCTGAGACCTGAGAGAGCTTGAAGCTGAGGGGAAACTCCGAGCT 2940
2881 GATTTCCCTAAGAACCTGAGAGCTTGAAGCTGAGGGGAAACTCCGAGAGCT 2940
2881 GATTTCCCTAAGAACCTGAGAGCTTGAAGCTGAGGGGAAACTCCGAGAGCT 2940
2941 GTCGAGACCCCGAGATGTCGTGCTGAGTACAGTACTTGTGAGCTTCCAGCTCCGA 3000
2941 GTCGAGACCCCGAGATGTCGTGCTGAGTACAGTACTTGTGAGCTTCCAGCTCCGA 3000
2941 GTCGAGACCCCGAGATGTCGTGCTGAGTACAGTACTTGTGAGCTTCCAGCTCCGA 3000
2941 GTCGAGACCCCGAGATGTCGTGCTGAGTACAGTACTTGTGAGCTTCCAGCTCCGA 3000
3001 GTTCACTTCTTATTAACCTGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3060
3001 GTTCACTTCTTATTAACCTGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3060
3001 GTTCACTTCTTATTAACCTGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3060
3001 GTTCACTTCTTATTAACCTGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3060
3061 GAGCCCATGAG 3120
3061 GAGCCCATGAG 3120
3121 ACCTCCAGCGAG 3180
3121 ACCTCCAGCGAG 3180
3121 ACCTCCAGCGAG 3180
3121 ACCTCCAGCGAG 3180
3181 CTTCATGTCGCGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3240
3181 CTTCATGTCGCGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3240
3241 CGTGGCGCTGAG 3300
3241 CGTGGCGCTGAG 3300
3241 CGTGGCGCTGAG 3300
3241 CGTGGCGCTGAG 3300
3301 CAACCTGGGCGAG 3360
3301 CAACCTGGGCGAG 3360
3301 CAACCTGGGCGAG 3360
3301 CAACCTGGGCGAG 3360
3361 CTAAGCGCGCTGTCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3420
3361 CTAAGCGCGCTGTCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3420
3421 CTAAGCGCGCTGTCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3480
3421 CTAAGCGCGCTGTCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3480
3421 CTAAGCGCGCTGTCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3480
3421 CTAAGCGCGCTGTCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3480
3481 AGAGCGGCGCTGCGAG 3540
3481 AGAGCGGCGCTGCGAG 3540
3481 AGAGCGGCGCTGCGAG 3540
3481 AGAGCGGCGCTGCGAG 3540
3541 TCATGAATGCTTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 3600
3541 TCATGAATGCTTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 3600
3541 TCATGAATGCTTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 3600
3541 TCATGAATGCTTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 3600
3601 GTGGCAAG 3660
3601 GTGGCAAG 3660
3601 GTGGCAAG 3660
3601 GTGGCAAG 3660
3661 GAGACACACCGGTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3720
3661 GAGACACACCGGTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3720
3661 GAGACACACCGGTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3720
3661 GAGACACACCGGTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3720
3721 CTGCTGGCGCGAGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3780
3721 CTGCTGGCGCGAGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3780
3721 CTGCTGGCGCGAGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3780
3721 CTGCTGGCGCGAGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3780
3781 GTTGTAAATAAAGTATTTAATA 3804
3781 GTTGTAAATAAAGTATTTAATA 3804
3781 GTTGTAAATAAAGTATTTAATA 3804
3781 GTTGTAAATAAAGTATTTAATA 3804

RESULT 2
ABK87218
ID ABK87218 standard; cdna: 3231 BP.
XX
AC
XX ABK87218;
XX

```
DT 24-SEP-2002 (first entry)
XX Human lipase, hormone-sensitive (LIPE) coding sequence.
DE Human lipase, hormone-sensitive (LIPE) coding sequence.
XX Human; lipase; hormone sensitive; LIPE; isogene; obesity;
KW male sterility; single nucleotide polymorphism; SNP;
XX Chromosome 19q13.1-q13.2; gene; ss.
OS Homo sapiens.

Location/Qualifiers
FT 1..3231
FT /tag- a
FT /product- "Human lipase, hormone sensitive (LIPE)
FT protein"
FT variation
FT replace (280,A)
FT /tag- b
FT /standard_name- "Single nucleotide polymorphism (SNP)"
FT /note- "Polymorphic site (PS)"
FT variation
FT replace (295,T)
FT /tag- c
FT /standard_name- "Single nucleotide polymorphism (SNP)"
FT /note- "Polymorphic site (PS)"
FT variation
FT replace (298,C)
FT /tag- d
FT /standard_name- "Single nucleotide polymorphism (SNP)"
FT /note- "Polymorphic site (PS)"
FT variation
FT replace (436,T)
FT /tag- e
FT /standard_name- "Single nucleotide polymorphism (SNP)"
FT /note- "Polymorphic site (PS)"
FT variation
FT replace (529,A)
FT /tag- f
FT /standard_name- "Single nucleotide polymorphism (SNP)"
FT /note- "Polymorphic site (PS)"
FT variation
FT replace (634,C)
FT /tag- g
FT /standard_name- "Single nucleotide polymorphism (SNP)"
FT /note- "Polymorphic site (PS)"
FT variation
FT replace (1210,T)
FT /tag- h
FT /standard_name- "Single nucleotide polymorphism (SNP)"
FT /note- "Polymorphic site (PS)"
FT variation
FT replace (1269,T)
FT /tag- i
FT /standard_name- "Single nucleotide polymorphism (SNP)"
FT /note- "Polymorphic site (PS)"
FT variation
FT replace (1626,T)
FT /tag- j
FT /standard_name- "Single nucleotide polymorphism (SNP)"
FT /note- "Polymorphic site (PS)"
FT variation
FT replace (1677,A)
FT /tag- k
FT /standard_name- "Single nucleotide polymorphism (SNP)"
FT /note- "Polymorphic site (PS)"
FT variation
FT replace (1831,T)
FT /tag- l
FT /standard_name- "Single nucleotide polymorphism (SNP)"
FT /note- "Polymorphic site (PS)"
FT variation
FT replace (1905,A)
FT /tag- m
FT /standard_name- "Single nucleotide polymorphism (SNP)"
FT /note- "Polymorphic site (PS)"
FT variation
FT replace (2368,A)
FT /tag- n
FT /standard_name- "Single nucleotide polymorphism (SNP)"
FT /note- "Polymorphic site (PS)"
FT variation
FT replace (2429,A)
FT /tag- o
FT /standard_name- "Single nucleotide polymorphism (SNP)"
FT /note- "Polymorphic site (PS)"
FT variation
FT replace (2453,A)
FT /tag- p

/standard_name- "Single nucleotide polymorphism (SNP)"
/note- "Polymorphic site (PS)"
replace (2508,T)
/tag- q
/standard_name- "Single nucleotide polymorphism (SNP)"
/note- "Polymorphic site (PS)"
replace (2812,A)
/tag- r
/standard_name- "Single nucleotide polymorphism (SNP)"
/note- "Polymorphic site (PS)"
replace (3000,C)
/tag- s
/standard_name- "Single nucleotide polymorphism (SNP)"
/note- "Polymorphic site (PS)"
replace (3171,T)
/tag- t
/standard_name- "Single nucleotide polymorphism (SNP)"
/note- "Polymorphic site (PS)"
replace (3201,A)
/tag- u
/standard_name- "Single nucleotide polymorphism (SNP)"
/note- "Polymorphic site (PS)"

W0200240502-A2.
XX
XX
XX 23-MAY-2002.
PD
XX
XX 16-NOV-2001; 2001WO-US43518.
PE
XX
XX 16-NOV-2000; 2000US-249302P.
PR
XX
XX (GENA-) GENAISSANCE PHARM INC.
PI
XX
XX Anastasio AE, Bentivegna SC, Chew A, Koshy B, Rounds E;
DR
DR WPI: 2002-519369/55.
DR P-PSDB; AA099253.
XX
XX Novel genetic variants of Lipase, Hormone-Sensitive isogenes, useful
PT for improving efficiency and reliability in drug development for
PT treating diseases associated with LIPE activity, e.g. obesity and male
PT sterility
PT
XX Claim 25; Fig 2; 142pp; English.
PS
XX
XX The present invention relates to a new polynucleotide comprising a
CC nucleotide sequence which comprises lipase, hormone sensitive (LIPE)
CC isogenes. The invention is useful in screening for drugs targeting LIPE
CC isogenes that are useful for treating obesity and male sterility. The
CC methods of the invention are useful for improving the efficiency and
CC reliability of several steps in the discovery and development of drugs
CC for treating diseases associated with LIPE activity. The polynucleotide
CC is useful in studying the expression and function of LIPE, and in
CC expressing LIPE protein for use in screening for candidate drugs to treat
CC diseases related to LIPE activity. It is also useful in studying the
CC effect of the variation on the biological activity of LIPE as well as on
CC the binding affinity of candidate drugs targeting LIPE for the treatment
CC of obesity and male sterility. The invention is useful for studying the
CC expression of LIPE isogenes in vivo, for in vivo screening and testing of
CC drugs targeted against LIPE protein, and for testing the efficacy of
CC therapeutic agents and compounds for treating obesity and male sterility
CC in a biological system. The present nucleic acid sequence represents the
CC coding sequence of the human LIPE gene located on chromosome
CC 19q13.1-q13.2. This sequence encodes the human LIPE protein of the
CC invention.
XX
SQ Sequence 3231 BP; 693 A; 1080 C; 906 G; 552 T; 0 other;
XX
XX Query Match 84.9%; Score 3227.8; DB 24; Length 3231;
XX Best Local Similarity 99.9%; Pred. No. 0;
XX Matches 3229; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY 278 ATGAGCCAGGTTCTAAGTCACTGCTAGGTGACACTGCAACCACTGACAGAG 337
```

```
|||||
Db 1 ATGAGCCAGGTTCTAAGTCAGTGTAGGTAGACTGCAACCTGAAACCAACACAGAG 60
QY 338 CCGATAACCCCGCTAGAGCTGGGCCGAAAGACACCCATAGCCCGCAAGATCGAG 397
Db 61 CCGATAACCCCGCTAGAGCTGGGCCGAAAGACACCCATAGCCCGCAAGATCGAG 120
QY 398 ACTCTGAGGATCCCAATACCAAGAGAGCTGCTTCAACCAAGAACCCCTACCCAG 457
Db 121 ACTCTGAGGATCCCAATACCAAGAGAGCTGCTTCAACCAAGAACCCCTACCCAG 180
QY 458 CAGGAGACCCCTGACAAACATGATGCTGAATCCCAAGAACCTAGAGCCCAAAAA 517
Db 181 CAGGAGACCCCTGACAAACATGATGCTGAATCCCAAGAACCTAGAGCCCAAAAA 240
QY 518 TCTGCTTCAACAAGAAATTTCTGCCCCCAGAGACCCGACCAACATCACTTAC 577
Db 241 TCTGCTTCAACAAGAAATTTCTGCCCCCAGAGACCCGACCAACATCACTTAC 300
QY 578 ATCCAAAGGCTGCTGCTCACTCAACAGAGAGCTGCTCCAGAGAGGAGCTGGCTAGA 637
Db 301 ATCCAAAGGCTGCTGCTCACTCAACAGAGAGCTGCTCCAGAGAGGAGCTGGCTAGA 360
QY 638 AAGAAATCTATACTCAACAGAGAGCCAGCATTTGAGACAAAGACATTTAGCCCAAGG 697
Db 361 AAGAAATCTATACTCAACAGAGAGCCAGCATTTGAGACAAAGACATTTAGCCCAAGG 420
QY 698 CCGGGCCAGAGAGAGCCACTCCAGCTCAACAAGAGCTAATACACCTGGGCCAG 757
Db 421 CCGGGCCAGAGAGAGCCACTCCAGCTCAACAAGAGCTAATACACCTGGGCCAG 480
QY 758 GCTAAACCTGAGCCAAAGAGAGCCATCTGCCCGACTGAATCTACATCCCAAGAGACA 817
Db 481 GCTAAACCTGAGCCAAAGAGAGCCATCTGCCCGACTGAATCTACATCCCAAGAGACA 540
QY 818 CCTGAACAGTCAGACAAGCAAGACGCCATCCAGGAGCCCAATCAAGCAAGGATCT 877
Db 541 CCTGAACAGTCAGACAAGCAAGACGCCATCCAGGAGCCCAATCAAGCAAGGATCT 600
QY 878 TTGACAGAGCTGGGATTTCTAACAAAATCTCAGAGACTATCCATACGCCATAGCCCTTA 937
Db 601 TTGACAGAGCTGGGATTTCTAACAAAATCTCAGAGACTATCCATACGCCATAGCCCTTA 660
QY 938 GAGTGAAGGCACTTTCTGAGTGGTGCAGATTCTGAGTCAGATAGATGTTGGATCA 997
Db 661 GAGTGAAGGCACTTTCTGAGTGGTGCAGATTCTGAGTCAGATAGATGTTGGATCA 720
QY 998 TCTTTCAGACACAGATTTCTCCAGCCAGATGGTGGAAATGGTGGCCAGGAGTGAACCTA 1057
Db 721 TCTTTCAGACACAGATTTCTCCAGCCAGATGGTGGAAATGGTGGCCAGGAGTGAACCTA 780
QY 1058 GGGTTCAAAGAAATCTGGTTTAAAGTATGTCAGAGATACAGTGGGAGCTGGCCCAT 1117
Db 781 GGGTTCAAAGAAATCTGGTTTAAAGTATGTCAGAGATACAGTGGGAGCTGGCCCAT 840
QY 1118 GAGAAAACAGTCTCGGATTCACAGACACTACAGAGATACAGGCTCAAGGCTCATCCAC 1177
Db 841 GAGAAAACAGTCTCGGATTCACAGACACTACAGAGATACAGGCTCAAGGCTCATCCAC 900
QY 1178 AACATGACCTGCGACAAATGACACAGTGGTGTGATCTTGGCGGAGCAACATAGCC 1237
Db 901 AACATGACCTGCGACAAATGACACAGTGGTGTGATCTTGGCGGAGCAACATAGCC 960
QY 1238 TTCTTCTCAGAGGCTCTGGGGAACGCGCCAGAGGCTGTACAGGCTTTTGGCGGT 1297
Db 961 TTCTTCTCAGAGGCTCTGGGGAACGCGCCAGAGGCTGTACAGGCTTTTGGCGGT 1020
QY 1298 GTACGGAGCAGAGCTGGGGCTGAGACCGGCTGGGCGCTGTGGGTGTGGCGCAC 1357
Db 1021 GTACGGAGCAGAGCTGGGGCTGAGACCGGCTGGGCGCTGTGGGTGTGGCGCAC 1080
QY 1358 CTCTTTGACCTGGAGCCAGAGACCGGCGCAACGGGATACCGAGCCTAATGTGACACAGCC 1417
|||||

Db 1081 CTCTTTGACCTGGAGCCAGAGACACCGGCGCAACGGGATACCGGCTAATGTGACACAGCC 1140
QY 1418 CGGTGCTGCTGGGACCTCTGACAAATCCGCTATGTGGCTTCCACCGCGGAGC 1477
Db 1141 CGGTGCTGCTGGGACCTCTGACAAATCCGCTATGTGGCTTCCACCGCGGAGC 1200
QY 1478 ATCTTCTTCCGACAGCACAACCTGGCGGAGCTGAGGAGCTACCTGGTGCCTCAC 1537
Db 1201 ATCTTCTTCCGACAGCACAACCTGGCGGAGCTGAGGAGCTTACTGTGGCTCAC 1260
QY 1538 CAGCTCGCGCTGTGCTACTAGCCAGCGCTGCTGTGTTACCAATGGCCGGGGTA 1597
Db 1261 CAGCTCGCGCTGTGCTACTAGCCAGCGCTGCTGTGTTACCAATGGCCGGGGTA 1320
QY 1598 CTCTTCTTTGAGGCGAGAGAGGGCTACCGGCTTCTCCGGGAATATGTACGCTG 1657
Db 1321 CTCTTCTTTGAGGCGAGAGAGGGCTACCGGCTTCTCCGGGAATATGTACGCTG 1380
QY 1658 CATTAAGGATCTTCTATAGGCGCTGCTGGGCTTCCAGTTCCAGGCTGCCATCCGGCCA 1717
Db 1381 CATTAAGGATCTTCTATAGGCGCTGCTGGGCTTCCAGTTCCAGGCTGCCATCCGGCCA 1440
QY 1718 TTCTGAGACCATCTCCATTTGGGCTGTCTTCCGGGAGACACTACAAACGAAAGAG 1777
Db 1441 TTCTGAGACCATCTCCATTTGGGCTGTCTTCCGGGAGACACTACAAACGAAAGAG 1500
QY 1778 ACAGGCTCAATGTGGCGGCAAGCTCTCTTACACAGCGCGCTTTGCCATGAGCC 1837
Db 1501 ACAGGCTCAATGTGGCGGCAAGCTCTCTTACACAGCGCGCTTTGCCATGAGCC 1560
QY 1838 GAGCTGGGCTGAGTTGAGCGGATCAACAGAACTGGAGCTGACCTTTCGAAA 1897
Db 1561 GAGCTGGGCTGAGTTGAGCGGATCAACAGAACTGGAGCTGACCTTTCGAAA 1620
QY 1898 GCGTTCTGGAACATACCGAGATGGAATGCTATGCTCTGGCCAAACATGGATGGCC 1957
Db 1621 GCGTTCTGGAACATACCGAGATGGAATGCTATGCTCTGGCCAAACATGGATGGCC 1680
QY 1958 ACCGTGAGGGTAAAGCCGCTCTGAGCTGGCACCCGAAAGCTTTGAAATGCCACTGACT 2017
Db 1681 ACCGTGAGGGTAAAGCCGCTCTGAGCTGGCACCCGAAAGCTTTGAAATGCCACTGACT 1740
QY 2018 GCCGACCCAGCTCAGCGGTACCATCTCACCCCACTGGGCCACACAGGCCCTGGGCC 2077
Db 1741 GCCGACCCAGCTCAGCGGTACCATCTCACCCCACTGGGCCACACAGGCCCTGGGCC 1800
QY 2078 GTCTCTGACGGCTATCTCTATGACCTGCTGAAGGACAGGACATGAGAGCTCAGC 2137
Db 1801 GTCTCTGACGGCTATCTCTATGACCTGCTGAAGGACAGGACATGAGAGCTCAGC 1860
QY 2138 AGCCTGATAAAGTCCAAAGGCGCAAGAGGCTGAGAGCTGTGGCCGCCCCACAGAGCA 2197
Db 1861 AGCCTGATAAAGTCCAAAGGCGCAAGAGGCTGAGAGCTGTGGCCGCCCCACAGAGCA 1920
QY 2198 CCGCGCTCGCGGCTCTGATATGATGACATCTCACGCGGCTGTGTTGGCCAGACCTCC 2257
Db 1921 CCGCGCTCGCGGCTCTGATATGATGACATCTCACGCGGCTGTGTTGGCCAGACCTCC 1980
QY 2258 AGATCCACGAGGCTCTACCTCAAGAGCTGGGCCAGAGAGCTGGGCGGCCCATCATCTCC 2317
Db 1981 AGATCCACGAGGCTCTACCTCAAGAGCTGGGCCAGAGAGCTGGGCGGCCCATCATCTCC 2040
QY 2318 ATGAGCTACCTCCGCGGCGGAGGCGGCTTCCCGGCTGGGCTGGAGAGATGCTTCTTC 2377
Db 2041 ATGAGCTACCTCCGCGGCGGAGGCGGCTTCCCGGCTGGGCTGGAGAGATGCTTCTTC 2100
QY 2378 GCGTACTGCTGGGCGCATCAAGCACTGGCCCTCTTGGCTCAAGAGGGAAGCATCTGC 2437
Db 2101 GCGTACTGCTGGGCGCATCAAGCACTGGCCCTCTTGGCTCAAGAGGGAAGCATCTGC 2160
QY 2438 CTTCGGGGGACAGTGCAGAGCGGGAACTCTGTCTTACCGTGTCTTGGGACAGACCC 2497
Db 2161 CTTCGGGGGACAGTGCAGAGCGGGAACTCTGTCTTACCGTGTCTTGGGACAGACCC 2220
|||||
```



```

FT      /note= "Polymorphic site 11 (PS11)"
FT      22811..22902
FT      /*tag= g
FT      /number= 3
FT      22903..23012
FT      /*tag= r
FT      /number= 3
FT      23013..23156
FT      /*tag= s
FT      /number= 4
FT      replace (23127,T)
FT      /*tag= t
FT      /standard_name= "Single nucleotide polymorphism (SNP)"
FT      /note= "Polymorphic site 12 (PS12)"
FT      23157..23340
FT      /*tag= u
FT      /number= 4
FT      23341..23527
FT      /*tag= v
FT      /number= 5
FT      replace (23362,A)
FT      /*tag= w
FT      /standard_name= "Single nucleotide polymorphism (SNP)"
FT      /note= "Polymorphic site 13 (PS13)"
FT      replace (23516,T)
FT      /*tag= x
FT      /standard_name= "Single nucleotide polymorphism (SNP)"
FT      /note= "Polymorphic site 14 (PS14)"
FT      23528..23664
FT      /*tag= y
FT      /number= 5
FT      replace (23608,A)
FT      /*tag= z
FT      /standard_name= "Single nucleotide polymorphism (SNP)"
FT      /note= "Polymorphic site 15 (PS15)"
FT      23665..23959
FT      /*tag= aa
FT      /number= 6
FT      replace (23727,A)
FT      /*tag= ab
FT      /standard_name= "Single nucleotide polymorphism (SNP)"
FT      /note= "Polymorphic site 16 (PS16)"
FT      23960..24744
FT      /*tag= ac
FT      /number= 6
FT      replace (24064,G)
FT      /*tag= ad
FT      /standard_name= "Single nucleotide polymorphism (SNP)"
FT      /note= "Polymorphic site 17 (PS17)"
FT      24745..24972
FT      /*tag= ae
FT      /number= 7
FT      24973..25571
FT      /*tag= af
FT      /number= 7
FT      25572..25748
FT      /*tag= ag
FT      /number= 8
FT      replace (25574,A)
FT      /*tag= ah
FT      /standard_name= "Single nucleotide polymorphism (SNP)"
FT      /note= "Polymorphic site 18 (PS18)"
FT      replace (25635,A)
FT      /*tag= ai
FT      /standard_name= "Single nucleotide polymorphism (SNP)"
FT      /note= "Polymorphic site 19 (PS19)"
FT      replace (25659,A)
FT      /*tag= aj
FT      /standard_name= "Single nucleotide polymorphism (SNP)"
FT      /note= "Polymorphic site 20 (PS20)"
FT      replace (25714,T)
FT      /*tag= ak
FT      /standard_name= "Single nucleotide polymorphism (SNP)"

```

```

FT      /note= "Polymorphic site 21 (PS21)"
FT      25749..28101
FT      /*tag= al
FT      /number= 8
FT      28102..28526
FT      /*tag= am
FT      /number= 9
FT      replace (28371,A)
FT      /*tag= an
FT      /standard_name= "Single nucleotide polymorphism (SNP)"
FT      /note= "Polymorphic site 22 (PS22)"
FT      28527..29057
FT      /*tag= ao
FT      /number= 9
FT      replace (28958,C)
FT      /*tag= ap
FT      /standard_name= "Single nucleotide polymorphism (SNP)"
FT      /note= "Polymorphic site 23 (PS23)"
FT      replace (29012,C)
FT      /*tag= ap
FT      /standard_name= "Single nucleotide polymorphism (SNP)"
FT      /note= "Polymorphic site 24 (PS24)"
FT      replace (29014,G)
FT      /*tag= ar
FT      /standard_name= "Single nucleotide polymorphism (SNP)"
FT      /note= "Polymorphic site 25 (PS25)"
FT      replace (29018,A)
FT      /*tag= as
FT      /standard_name= "Single nucleotide polymorphism (SNP)"
FT      /note= "Polymorphic site 26 (PS26)"
FT      29058..29321
FT      /*tag= at
FT      /number= 10
FT      replace (29090,C)
FT      /*tag= au
FT      /standard_name= "Single nucleotide polymorphism (SNP)"
FT      /note= "Polymorphic site 27 (PS27)"
FT      replace (29261,T)
FT      /*tag= av
FT      /standard_name= "Single nucleotide polymorphism (SNP)"
FT      /note= "Polymorphic site 28 (PS28)"
FT      replace (29291,A)
FT      /*tag= aw
FT      /standard_name= "Single nucleotide polymorphism (SNP)"
FT      /note= "Polymorphic site 29 (PS29)"
FT      replace (29303,A)
FT      /*tag= ax
FT      /standard_name= "Single nucleotide polymorphism (SNP)"
FT      /note= "Polymorphic site 30 (PS30)"
FT      replace (29640,C)
FT      /*tag= ay
FT      /standard_name= "Single nucleotide polymorphism (SNP)"
FT      /note= "Polymorphic site 31 (PS31)"

```

Query Match 30.3%; Score 1152.8; DB 24; Length 30709;
Best Local Similarity 99.2%; Pred. No. 2.1e-239;
Matches 1151; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

```

QY      1 CTTCTTGTAGAGAGTGTAGACACATAGCCCCCTCTATTCTATCTATCTCCACCAAG 60
DB      CTTCTTGTAGAGAGTGTAGACACATAGCCCTCTATTCTATCTCCACCAAG 3866
QY      61 AAAGAGGCACAGAGTTCATTAGTGTGGGGCCAGCTGTATGCGCCAACTGCCAGCTG 120
DB      AAAGAGGCACAGAGTTCATTAGTGTGGGGCCAGCTGTATGCGCCAACTGCCAGCTG 3926
QY      121 CCTTAAAGAGAACGCCAGTGTAGATGAGATGAACCCAGAGAGAGTGCATCA 180
DB      CCTTAAAGAGAACGCCAGTGTAGATGAGATGAACCCAGAGAGAGTGCATCA 3986
QY      3927 CCTTAAAGAGAACGCCAGTGTAGATGAGATGAACCCAGAGAGAGTGCATCA 3986
DB      181 TGAGGAATCAATGAGAGATCTGTGAAGAGAGAGGGCTGGTGGAGCCAGAGAGATAGA 240
DB      3987 TGAGGAATCAATGAGAGATCTGTGAAGAGAGAGGGCTGGTGGAGCCAGAGAGATAGA 4046

```

```

OY 241 ACCTGGAAGATCAATATCTCCGCTGAGGGAATTAACATGAGCAAGTTCTAAGTCACT 300
DB 4047 ACCTGGAAGATCAATATCTCCGCTGAGGGAATTAACATGAGCAAGTTCTAAGTCACT 4106
OY 301 GCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
DB 4107 GCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4166
OY 361 GCCAGAGAAAGACACCATAGCCAGCAGAAATGAAAGCTGTGAGGATTCATACCA 420
DB 4167 GCCAGAGAAAGACACCATAGCCAGCAGAAATGAAAGCTGTGAGGATTCATACCA 4226
OY 421 ACAGAGCCCTGCTTAAACCAAGACCCCTCACCCAGAGAGACCCCTGCACCAACATGA 480
DB 4227 ACAGAGCCCTGCTTAAACCAAGACCCCTCACCCAGAGAGACCCCTGCACCAACATGA 4286
OY 481 TGCTGAATCCCAAGAGAACTAGAGCCCAACAAAATCTGCTTACCAAGAGAAATTTCT 540
DB 4287 TGCTGAATCCCAAGAGAACTAGAGCCCAACAAAATCTGCTTACCAAGAGAAATTTCT 4346
OY 541 TGCCCCACAGAGAGCCGACACACATCACTTACATTCACAAAGGGTGTCTCACTCA 600
DB 4347 TGCCCCACAGAGAGCCGACACACATCACTTACATTCACAAAGGGTGTCTCACTCA 4406
OY 601 ACAGAGAGCTGCTCCACAGAGAGAGCTGGGCTAGAGAAAAGATCTATTAATCAACAGAG 660
DB 4407 ACAGAGAGCTGCTCCACAGAGAGAGCTGGGCTAGAGAAAAGATCTATTAATCAACAGAG 4466
OY 661 GCCAGCATTTGAGACAAAACATGTAGCCAGCAGGCTGGGCTGGGCTGGGCTGGGCTGGG 720
DB 4467 GCCAGCATTTGAGACAAAACATGTAGCCAGCAGGCTGGGCTGGGCTGGGCTGGGCTGGG 4526
OY 721 AGCTCAACAAAGAGTGAATCAACACTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG 780
DB 4527 AGCTCAACAAAGAGTGAATCAACACTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG 4586
OY 781 GCCATCTGCCCCGAGTGAATCTACATCCCAAGAGACACCTGAACAGTCAAGACCAAC 840
DB 4587 GCCATCTGCCCCGAGTGAATCTACATCCCAAGAGACACCTGAACAGTCAAGACCAAC 4646
OY 841 AAGCGAGTCCAGGAGGCAATCCAAACAGAGAGCTTTGACAGAGCTGGGCTGGGCTGGG 900
DB 4647 AAGCGAGTCCAGGAGGCAATCCAAACAGAGAGCTTTGACAGAGCTGGGCTGGGCTGGG 4706
OY 901 AAAACTTCAGAGACTATCATACAGCGATCAAGCCCTAGAGTGAAGGCACTTCTGAGTG 960
DB 4707 AAAACTTCAGAGACTATCATACAGCGATCAAGCCCTAGAGTGAAGGCACTTCTGAGTG 4766
OY 961 GGTGCGAGATTTGAGTCAAGATCAGATGAGGATCATCTTCAGACAGAGATTTCTCAGC 1020
DB 4767 GGTGCGAGATTTGAGTCAAGATCAGATGAGGATCATCTTCAGACAGAGATTTCTCAGC 4826
OY 1021 CAGCATGGTGGAAATGGTGGCCAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1080
DB 4827 CAGCATGGTGGAAATGGTGGCCAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 4886
OY 1081 TAAAGTGAATGTCAGAGATCACTGGAGCTCCGACATGAGAAAAACAGTGTCTGGAATCA 1140
DB 4887 TAAAGTGAATGTCAGAGATCACTGGAGCTCCGACATGAGAAAAACAGTGTCTGGAATCA 4946
OY 1141 CAGACACTACAGAGATACAG 1160
DB 4947 CAGACACTACAGAGATACAG 4966

```

```

RESULT 4
ABQ25066
ID ABQ25066 standard; DNA; 709 BP.
XX
AC ABQ25066;
XX
DT 12-JUL-2002 (first entry)

```

```

XX DE Oligonucleotide for detecting cytosine methylation SEQ ID NO 11657.
XX XX Human: cytosine methylation; 5'-CPG-3'; uracil; cytosine; diagnosis;
XX KM drug; side effect; cancer; central nervous system; cardiovascular;
XX KM gastrointestinal; respiratory system; single nucleotide polymorphism;
XX KM SNP; cell differentiation; ds.
OS Homo sapiens.
PN WO200218632-A2.
XX
PD 07-MAR-2002.
XX
PF 01-SEP-2001; 2001WO-EPI0074.
XX
PR 01-SEP-2000; 2000DE-1043826.
XX
PR 05-SEP-2000; 2000DE-1044543.
XX
PA (EPIC-) EPIGENOMICS AG.
XX
PI Olek A., Piepenbrock C., Berlin K., Guetig D;
XX
XX WPI: 2002-371829/40.
XX
XX Determining the degree of cytosine methylation in genomic DNA, useful
XX PT for diagnosis and prognosis, comprises selective hybridization of
XX PT amplicons from chemically treated DNA
XX
XX Claim 12; 56pp + Sequence Listing; 56pp; German.
XX
XX This invention describes a novel method for determining the degree of
XX CC methylation of a particular cytosine in a motif 5'-CPG-3', present in a
XX CC genomic sample of DNA. The sample is treated chemically to convert
XX CC cytosine (C) but not methylated C, to uracil, then part of the genomic
XX CC DNA that contains the target C is amplified to form a labeled amplicon.
XX CC The amplicon is hybridised to two classes, each with at least one
XX CC member, of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers
XX CC and the degree of hybridisation to both classes is determined from the
XX CC label on the amplicon. From the ratio of labels hybridised to the two
XX CC classes of oligomers, the degree of methylation is calculated. The method
XX CC is used: (i) for diagnosis and/or prognosis of side effects of
XX CC therapeutic drugs and of a wide range of diseases, e.g. cancer, disorders
XX CC of the central nervous, cardiovascular, gastrointestinal and respiratory
XX CC systems etc., particularly by detecting mutations or single nucleotide
XX CC polymorphisms (SNP's); and (ii) for differentiation of cell or tissue
XX CC types and for investigating cell differentiation. The method allows the
XX CC methylation status of many C residues to be determined simultaneously.
XX CC ABQ13410-ABQ34121 represent genomic DNA sequences used to illustrate the
XX CC method for determining the degree of cytosine methylation described in
XX CC the disclosure of the invention.
SO Sequence 709 BP; 97 A; 74 C; 238 G; 300 T; 0 other;

```

```

Query Match 9.2%; Score 348.4; DB 24; Length 709;
Best Local Similarity 79.5%; Pred. No. 8.9e-66;
Matches 450; Conservative 0; Mismatches 111; Indels 5; Gaps 3;

```

```

OY 3244 GCGGTGCGCGTGGAGCCCATGCTGAGACGACGCTGCGGCGGCGAGTGGCA 3303
DB 119 GCGGTGCGCGTGGAGTATTTATGTTGAGATTCGTTATGTTGCGGCGGCGAGTGGCA 178
OY 3304 CTTGGGCGAGCGGAGCGTGGCGGCTGGTGGAGAGCTGCGCGAGCGCTTCTGACCT 3363
DB 179 TTTGGTTAGTGGTGGAGCGTGGCGGCTGGTGGAGAGATTTGCTGACGTTTATTTT 238
OY 3364 AGCGGCGTGTGCGCGAGAGCGGCGAGCGGCGAGAGCTGTGCGTGGAGCGCATCGCT 3423
DB 239 AGCGGCGTGTGCGCGAGAGCGGCGTGTGAGTGTGAGTGTGCGTGGAGCGCATTCGTTT 298
OY 3424 CGTCTCACTCTCCCGCGGAGCGGCGGCGGAGAGCGGCGGCGTGGGCGGAGTGA 3483
DB 299 CGTTTATTTTTCGTCGAGTGGGTGAGCGGCGGAGAGCGGCGGCGGCGGAGTGA 358

```


GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 3, 2003, 07:06:16 : Search time 499 Seconds
(without alignments)
17167.538 Million cell updates/sec

Title: US-09-915-814-3

Perfect score: 3804

Sequence: 1 ctctctgtaagaagagtgcta.....taataaaagatttaatta 3804

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 2390332

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

N.Geneseq_101002:*

- 1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
- 2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
- 3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
- 4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
- 5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
- 6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
- 7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
- 8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
- 9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
- 10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
- 11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
- 12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
- 13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
- 14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
- 15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
- 16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
- 17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
- 18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
- 19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
- 20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
- 21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
- 22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
- 23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
- 24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	60	1.6	60	24	ABN32626
2	51	1.3	51	22	ABL33801
3	51	1.3	51	22	ABL33804
4	50.6	1.3	65	24	ABN28875
5	39	1.0	50	22	ABL33802
6	39	1.0	50	22	ABL33803
7	37.2	1.0	65	24	ABN53647
8	30.8	0.8	65	24	ABN28121
9	29.2	0.8	65	24	ABN52192

10	29	0.8	88	23	AA548737	Pseudomonas aerugi	
C	11	27.6	0.7	89	19	AAV68382	Clone #8 fragment
C	12	27.6	0.7	60	24	ABN45730	Human spliced tran
C	13	27	0.7	60	24	ABN32825	Human spliced tran
C	14	27	0.7	72	20	AAx88105	Antifreeze protein
C	15	26.8	0.7	99	17	AAx27118	Hepatitis E virus
C	16	26.8	0.7	99	18	AAx96968	Hepatitis E virus
C	17	26.6	0.7	65	21	AAx88283	SLP111 nucleotide
C	18	26.6	0.7	65	22	AAx74335	Repetitive protein
C	19	26.6	0.7	65	22	AAx23371	SLP111 gene fragme
C	20	26.4	0.7	100	22	AAx19814	Human brain expres
C	21	26.4	0.7	100	22	AAx45837	Human bone marrow
C	22	26.4	0.7	100	24	ABx20120	Human genome-deriv
C	23	26.2	0.7	86	24	AAx98931	Human prostate can
C	24	26.2	0.7	87	16	AAx22575	Human gene signatu
C	25	26	0.7	90	24	ABx36582	HCV DNA encoding H
C	26	26	0.7	92	21	AAx11989	Human secreted pro
C	27	26	0.7	94	19	AAx38341	Manic-depressive 1
C	28	25.8	0.7	99	21	AAx36736	Muv and Mv fusion
C	29	25.6	0.7	64	20	AAx88104	Synthetic antifree
C	30	25.6	0.7	65	24	ABx57988	Mouse spliced tran
C	31	25.6	0.7	66	12	AAQ12354	Europium-labelled
C	32	25.6	0.7	70	21	AAx38176	Primer used in ste
C	33	25.6	0.7	99	14	AAQ36997	Proinsulin oligonu
C	34	25.6	0.7	99	21	AAx98715	Human colon cancer
C	35	25.6	0.7	100	24	ABx74850	Corn tassal-derive
C	36	25.4	0.7	96	21	AAx57144	Exemplary quadripl
C	37	25.4	0.7	100	17	AAx15244	Partial coding seq
C	38	25.2	0.7	61	18	AAx84492	SLP111 gene sectio
C	39	25.2	0.7	61	21	AAx82822	SLP111 nucleotide
C	40	25.2	0.7	61	22	AAx74334	Repetitive protein
C	41	25.2	0.7	61	22	AAx23370	SLP111 gene fragme
C	42	25.2	0.7	65	19	AAx41235	Mouse embryonic ce
C	43	25.2	0.7	68	22	AAx32956	DNA encoding CARD
C	44	25.2	0.7	100	22	AAx41935	Human CLASP-3 geno
C	45	25	0.7	66	17	AAx97798	Nucleic acid inh

ALIGNMENTS

RESULT 1
ABN32626
ID ABN32626 standard; DNA: 60 BP.
XX
AC ABN32626;
XX
DT 15-JUL-2002 (first entry)
XX
DE Human spliced transcript detection oligonucleotide SEQ ID NO:5374.
XX
KW Human; mouse; rat; splice transcript; detection; RNA transcript;
KW splice variant; transcriptome; oligonucleotide library; ss.
XX
OS Homo sapiens.
XX
PN WO200210449-A2.
XX
PD 07-FEB-2002.
XX
PF 20-JUL-2001; 2001WO-IB01903.
XX
PR 28-JUL-2000; 2000US-221607P.
XX
PR 02-MAY-2001; 2001US-287724P.
XX
PA (COMP-) COMPUGEN INC.
XX
PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
XX
DR WPI; 2002-257383/30.
XX
PT New oligonucleotide libraries comprising oligonucleotides which selectively hybridize to mRNAs transcribed from a transcription unit of

PT a genome, useful for detecting tissue-, pathology-, and
PT developmental-specific genes

XX Example 1; SEQ ID 5374; 47pp; English.

CC The present invention describes oligonucleotide libraries for detecting
CC messenger RNAs that populate a (sub-)transcriptome, where the
CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC transcription units that populate a genome. The library comprises
CC several oligonucleotides, each capable of hybridizing selectively to a
CC set of messenger RNAs transcribed from a given transcription unit of
CC the genome, which encodes one or more messenger RNA splice variants.
CC The oligonucleotide libraries are useful for detecting mRNAs from a
CC biological sample, in expression profiling studies, in qualitatively or
CC quantitatively characterizing the corresponding transcriptome, and in
CC detecting RNA transcripts and splice variants of human or animal
CC transcripts. The libraries may also be used as specialized mini
CC libraries to detect transcripts of a sub-transcriptome under a
CC detection of tissue- and pathology-specific genes such as those genes
CC only expressed in specific tissue under a specific pathological
CC condition; to detect developmental specific genes; and to detect RNA
CC transcripts and splice variants of a transcriptome of a patient suffering
CC from a particular disorder. ABN27253 to ABN55589 represent
CC oligonucleotide sequences from rats, humans and mice, which are used in
CC the exemplification of the present invention.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 60 BP; 19 A; 16 C; 16 G; 9 T; 0 other;

Query Match 1.6%; Score 60; DB 24; Length 60;

Best Local Similarity 100.0%; Pred. No. 0.0012;

Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2621 AAGTGTGTCAGCCGCTATGCTGTGCAAGAGAGAGACACACACACACACAGAA 2680
DB 1 AAGTGTGTCAGCCGCTATGCTGTGCAAGAGAGAGACACACACACACACAGAA 60

RESULT 2
AAL33801

ID AAL33801 standard; DNA; 51 BP.

XX AAL33801;

DT 24-JAN-2002 (first entry)

DE Human SNP oligonucleotide #7009.

XX Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
KW neuroprotective; antimicrobial; gene therapy; vaccine; amylose; cancer;
KW amyloid protein; angiotensin; apoptosis related protein; cadherin;
KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
KW complement related protein; cytochrome; kinesin; cytokine; interferon;
KW interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW multifactorial disease; autoimmune disease; infection;
KW nervous system disease; ss.

KW Homo sapiens.

OS Homo sapiens.

PN WO200147944-A2.

PD 05-JUL-2001.

PF 28-DEC-2000; 2000WO-US35498.

PR 28-DEC-1999; 99US-0173419.

PA 27-DEC-2000; 2000US-0173419.

XX (CURA-) CURAGEN CORP.

PI Shinkets RA, Leach M;

XX WPI; 2001-465210/50.

DR Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
XX oncogenes and histones, useful for diagnosing and treating, e.g.
PT cancer, autoimmune diseases and infections

PS Claim 1; Page 3391; 4143pp; English.

CC The present invention relates to oligonucleotides encoding polymorphic
CC variants of proteins related to amylases, amyloid proteins, angiotensin,
CC apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,
CC histones, kinases, colony stimulating factors, complement related
CC proteins, cytochromes, kinesins, cytokines, interferons, interleukins,
CC G-protein coupled receptors and thioesterases. The present sequence is
CC one such oligonucleotide. The oligonucleotides and the peptides encoded
CC by them may be used in the prevention, diagnosis and treatment of
CC diseases associated with inappropriate expression of the proteins listed
CC above. Disorders that may be prevented, diagnosed and/or treated include
CC multifactorial diseases with a genetic component, such as autoimmune
CC diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,
CC systemic lupus erythematosus and Grave's disease), inflammation, cancer
CC (e.g. cancers of the bladder, brain, breast, colon and kidney,
CC leukaemia), diseases of the nervous system and an infection of pathogenic
CC organisms.

SQ Sequence 51 BP; 6 A; 24 C; 11 G; 10 T; 0 other;

Query Match 1.3%; Score 51; DB 22; Length 51;

Best Local Similarity 100.0%; Pred. No. 0.0099;

Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1516 GGCCTACCTGGCTGCTCCACACAGCTCCGCGCTGCTACTACGCCCA 1566
DB 1 GGCCTACCTGGCTGCTCCACACAGCTCCGCGCTGCTACTACGCCCA 51

RESULT 3
AAL33804

ID AAL33804 standard; DNA; 51 BP.

XX AAL33804;

DT 24-JAN-2002 (first entry)

DE Human SNP oligonucleotide #7012.

XX Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
KW neuroprotective; antimicrobial; gene therapy; vaccine; amylose; cancer;
KW amyloid protein; angiotensin; apoptosis related protein; cadherin;
KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
KW complement related protein; cytochrome; kinesin; cytokine; interferon;
KW interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW multifactorial disease; autoimmune disease; infection;
KW nervous system disease; ss.

KW Homo sapiens.

OS Homo sapiens.

PN WO200147944-A2.

PD 05-JUL-2001.

PF 28-DEC-2000; 2000WO-US35498.

PR 28-DEC-1999; 99US-0173419.

PA 27-DEC-2000; 2000US-0173419.

XX (CURA-) CURAGEN CORP.

XX Shinkets RA, Leach M;

XX WPI; 2001-465210/50.

CC The present invention describes oligonucleotide libraries for detecting

PI Shinkets RA, Leach M;
VY

XX

DR WPI; 2001-465210/50.
 XX Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
 PT oncogenes and histones, useful for diagnosing and treating, e.g.
 PT cancer, autoimmune diseases and infections -
 PS Claim 1; Page 3392; 4143pp; English.
 CC The present invention relates to oligonucleotides encoding polymorphic
 CC variants of proteins related to amylases, amyloid proteins, angiotensin,
 CC apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,
 CC histones, kinases, colony stimulating factors, complement related
 CC proteins, cytochromes, kinesins, cytokines, interferons, interleukins,
 CC G-protein coupled receptors and thioesterases. The present sequence is
 CC one such oligonucleotide. The oligonucleotides and the peptides encoded
 CC by them may be used in the prevention, diagnosis and treatment of
 CC diseases associated with inappropriate expression of the proteins listed
 CC above. Disorders that may be prevented, diagnosed and/or treated include
 CC multifactorial diseases with a genetic component, such as autoimmune
 CC diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,
 CC systemic lupus erythematosus and Grave's disease), inflammation, cancer
 CC (e.g. cancers of the bladder, brain, breast, colon and kidney,
 CC leukaemia), diseases of the nervous system and an infection of pathogenic
 CC organisms.
 CC
 SQ Sequence 50 BP; 7 A; 16 C; 15 G; 12 T; 0 other;
 Query Match 1.0%; Score 39; DB 22; Length 50;
 Best Local Similarity 98.0%; Pred. No. 39;
 Matches 50; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
 QY 1596 TACTCTCTTTGAGGCGACGAGGGCTCACCCTCTCTCTCGGAGT 1646
 Db 1 TACTCTCTTTGAGGCGACGAGGGCTCACCCTCTCTCTCGGAGT 50
 RESULT 6
 AAL33803
 ID AAL33803 standard; DNA; 50 BP.
 XX
 AC AAL33803;
 XX
 DT 24-JAN-2002 (first entry)
 XX
 DE Human SNP oligonucleotide #7011.
 XX
 KW Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
 KW neuroprotective; antimicrobial; gene therapy; vaccine; amylose; cancer;
 KW amyloid protein; angiotensin; apoptosis related protein; cadherin;
 KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
 KW complement related protein; cytochrome; kinesin; cytokine; interferon;
 KW interleukin; G-protein coupled receptor; thioesterase; inflammation;
 KW multifactorial disease; autoimmune disease; infection;
 KW nervous system disease; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200147944-A2.
 XX
 PD 05-JUL-2001.
 XX
 PF 28-DEC-2000; 2000WO-US35498.
 XX
 PR 28-DEC-1999; 99US-0173419.
 PR 27-DEC-2000; 2000US-0173419.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Shimkets RA, Leach M;
 DR WPI; 2001-465210/50.
 PT Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,

PT oncogenes and histones, useful for diagnosing and treating, e.g.
 PT cancer, autoimmune diseases and infections -
 PS Claim 1; Page 3392; 4143pp; English.
 CC The present invention relates to oligonucleotides encoding polymorphic
 CC variants of proteins related to amylases, amyloid proteins, angiotensin,
 CC apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,
 CC histones, kinases, colony stimulating factors, complement related
 CC proteins, cytochromes, kinesins, cytokines, interferons, interleukins,
 CC G-protein coupled receptors and thioesterases. The present sequence is
 CC one such oligonucleotide. The oligonucleotides and the peptides encoded
 CC by them may be used in the prevention, diagnosis and treatment of
 CC diseases associated with inappropriate expression of the proteins listed
 CC above. Disorders that may be prevented, diagnosed and/or treated include
 CC multifactorial diseases with a genetic component, such as autoimmune
 CC diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,
 CC systemic lupus erythematosus and Grave's disease), inflammation, cancer
 CC (e.g. cancers of the bladder, brain, breast, colon and kidney,
 CC leukaemia), diseases of the nervous system and an infection of pathogenic
 CC organisms.
 CC
 SQ Sequence 50 BP; 10 A; 16 C; 15 G; 9 T; 0 other;
 Query Match 1.0%; Score 39; DB 22; Length 50;
 Best Local Similarity 98.0%; Pred. No. 39;
 Matches 50; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
 QY 1612 GCACGAGGGGCTCACCCTCTCTCGGAGATGTCACGTCGCAATA 1662
 Db 1 GCACGAGGGGCTCACCCTCTCTCGGAGATGTCACGTCGCAATA 50
 RESULT 7
 ABN53647
 ID ABN53647 standard; DNA; 65 BP.
 XX
 AC ABN53647;
 XX
 DT 15-JUL-2002 (first entry)
 XX
 DE Mouse spliced transcript detection oligonucleotide SEQ ID NO:26395.
 XX
 KW Human; mouse; rat; splice transcript; detection; RNA transcript;
 KW splice variant; transcriptome; oligonucleotide library; ss.
 XX
 OS Mus musculus.
 XX
 PN WO200210449-A2.
 XX
 PD 07-FEB-2002.
 XX
 PF 20-JUL-2001; 2001WO-IB01903.
 XX
 PR 28-JUL-2000; 2000US-221607P.
 PR 02-MAY-2001; 2001US-287724P.
 XX
 PA (COMP-) COMPUGEN INC.
 XX
 PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
 DR WPI; 2002-257383/30.
 XX
 PT New oligonucleotide libraries comprising oligonucleotides which
 PT selectively hybridize to mRNAs transcribed from a transcription unit of
 PT a genome, useful for detecting tissue-, pathology-, and
 PT developmental-specific genes -
 PS Example 1; SEQ ID 26395; 47pp; English.
 XX
 CC The present invention describes oligonucleotide libraries for detecting
 CC messenger RNAs that populate a (sub-)transcriptome, where the
 CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 3, 2003, 04:03:01 ; Search time 3324 Seconds
(without alignments)
18534.185 Million cell updates/sec

Title: US-09-915-814-3
Perfect score: 3804
Sequence: 1 cctcttgtaagagagtgctca.....taataaagaatttaatta 3804

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estlov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_huv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_trod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1792.2	47.1	1809	11	BC029961 Homo sapi
2	1213	31.9	2760	11	AK019846 Mus muscu
3	921	24.2	1084	13	BM563641 AGENCOURT
4	808.6	21.3	826	13	BI826568 603077016
5	787	20.7	872	13	BI827841 603074047
6	784	20.6	905	13	BI827706 603074127

7	782.6	20.6	870	13	BI827559
8	764.8	20.1	871	13 <td>BI561175</td>	BI561175
9	757	19.9	787	13 <td>BI830913</td>	BI830913
10	753.2	19.8	788	13 <td>BI520246</td>	BI520246
11	726	19.1	773	13 <td>BI520160</td>	BI520160
12	711.2	18.7	781	13 <td>BI832284</td>	BI832284
13	697	18.3	767	13 <td>AI337272</td>	AI337272
14	668.8	17.6	1367	11 <td>BC029301</td>	BC029301
15	652	17.1	663	13 <td>BI826036</td>	BI826036
16	645	17.0	645	13 <td>BI825158</td>	BI825158
17	623	16.4	625	13 <td>BM310576</td>	BM310576
18	619	16.3	668	14 <td>BQ028786</td>	BQ028786
19	613.8	16.1	736	9 <td>AI953697</td>	AI953697
20	611	16.1	691	13 <td>BI855956</td>	BI855956
21	610	16.0	641	13 <td>BI818150</td>	BI818150
22	568	14.9	569	10 <td>AM269980</td>	AM269980
23	567.2	14.9	588	14 <td>BQ807836</td>	BQ807836
24	558	14.7	985	12 <td>BG023944</td>	BG023944
25	556.2	14.6	653	9 <td>AI339344</td>	AI339344
26	555.8	14.6	601	9 <td>AI582719</td>	AI582719
27	548.2	14.4	555	9 <td>AI798896</td>	AI798896
28	538.6	14.2	635	9 <td>AA601541</td>	AA601541
29	534	14.0	534	12 <td>BE855571</td>	BE855571
30	521.4	13.7	741	12 <td>BG872368</td>	BG872368
31	520.4	13.7	524	10 <td>AM195614</td>	AM195614
32	510.4	13.4	892	14 <td>BQ953768</td>	BQ953768
33	503.2	13.2	669	9 <td>AA861626</td>	AA861626
34	500.6	13.2	1036	13 <td>BM560162</td>	BM560162
35	499.6	13.1	557	14 <td>BM714935</td>	BM714935
36	498.8	13.1	563	14 <td>BM676647</td>	BM676647
37	496.6	13.1	572	9 <td>AA639729</td>	AA639729
38	496	13.0	539	10 <td>AM027306</td>	AM027306
39	492.2	12.9	643	12 <td>BF222934</td>	BF222934
40	479	12.6	952	12 <td>BF537918</td>	BF537918
41	477.4	12.5	882	12 <td>BG868403</td>	BG868403
42	477.2	12.5	913	13 <td>BI654202</td>	BI654202
43	473.2	12.4	600	12 <td>BG803367</td>	BG803367
44	473	12.4	588	12 <td>BG835117</td>	BG835117
45	469	12.3	469	10 <td>AM136527</td>	AM136527

ALIGNMENTS

RESULT 1
BC029961
LOCUS 1809 bp mRNA linear HTC 06-MAY-2002
DEFINITION Homo sapiens, clone IMAGE:5171623, mRNA.
ACCESSION BC029961
VERSION BC029961.1 GI:20455827
KEYWORDS
SOURCE
ORGANISM Homo sapiens.
Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (06-MAY-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK
COMMENT NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: amg@bcm.tmc.edu

cdna Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: L14M12764 row: P column: 08
 High quality sequence start: 24
 High quality sequence stop: 713.

FEATURES

source

```

1.1084
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5744119"
/clone_lib="NIH_MGC_119"
/tissue_type="medulla"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMV-Sport6; Site:1: NotI;
Site:2: EcoRV (destroyed); RNA source normal medulla from
anonymous male age 27. Library is oligo-dT primed and
directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.3 kb, insert size range
0.9-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(invitrogen). Research Genetics tracking code 013. Note:
this is a NIH_MGC Library."
BASE COUNT      328 a      301 c      277 g      176 t      2 others
ORIGIN

```

Query Match 24.2%; Score 921; DB 13; Length 1084;
 Best Local Similarity 95.3%; Pred. No. 1.9e-198;
 Matches 1013; Conservative 0; Mismatches 42; Indels 8; Gaps 6;

```

OY 12 AGAGGTGCTAGTGCACATAGCCCTCTCTATTCCTAATCCCTCCCAAGAAAGAGGCACA 71
DB 24 AAGAGAGCTAGGACCTAGCTCTCTCTATTCCTAATCCCTCCCAAGAAAGAGGCACA 83
OY 72 GAGTTCATTACTAGTGGGGCCAGCTGTGATGCGCCCACTGCGCACTGCTTTAAAGAG 131
DB 84 GAGTTCATTACTAGTGGGGCCAGCTGTGATGCGCCCACTGCGCACTGCTTTAAAGAG 143
OY 132 AAGACAGTGTAGTGTGATGTAACCAAGAGAGAGTGCATGATGAGCAATCAA 191
DB 144 AAGACAGTGTAGTGTGATGTAACCAAGAGAGAGTGCATGATGAGCAATCAA 203
OY 192 TGAGAGATCTGTGAAGAGAGAGGCTGGTGGAGAGCCCAAGAGATAGAACTGGAAGAT 251
DB 204 TGAGAGATCTGTGAAGAGAGAGGCTGGTGGAGAGCCCAAGAGATAGAACTGGAAGAT 263
OY 252 CAATATCTCCCTGAGAGGAATTAACATGAGGACGAGTCTAAGTCAAGTCTAGGTCAG 311
DB 264 CAATATCTCCCTGAGAGGAATTAACATGAGGACGAGTCTAAGTCAAGTCTAGGTCAG 323
OY 312 ACTGGCAACTGACACACACAGAGGCTATTAACCCCGCTAGAGCCTGGGCCAGAAAGA 371
DB 324 ACTGGCAACTGACACACACAGAGGCTATTAACCCCGCTAGAGCCTGGGCCAGAAAGA 383
OY 372 CACCATATGCCAGGACGAGATGAGAGACTCTGCAGGATCCCAATACCAAGAGAACCCG 431
DB 384 CACCATATGCCAGGACGAGATGAGAGACTCTGCAGGATCCCAATACCAAGAGAACCCG 443
OY 432 CTTCAACCAAGACCCCTCACCCAGAGAGACCCCTGCACAAACATGATGTCGATCC 491
DB 444 CTTCAACCAAGACCCCTCACCCAGAGAGACCCCTGCACAAACATGATGTCGATCC 503
OY 492 AGAAGAACTTAGAGCCCAACAAATCTGCTTCACAGAGAGATTTCTGGCCCAAGA 551
DB 504 AGAAGAACTTAGAGCCCAACAAATCTGCTTCACAGAGAGATTTCTGGCCCAAGA 563
OY 552 AGCCGCAACAGAGATCCTTACATCCAAAGGCTGCTCACTCACTCAAGAGAGCG 611
DB 564 AGCCGCAACAGAGATCCTTACATCCAAAGGCTGCTCACTCACTCAAGAGAGCG 623
OY 612 CCTCCAGAGGACCTGGGCTAGGAAAAAGAAATCTAATCAACAGAGACGACATTTGA 671

```

```

DB 624 CTTCCAGAGAGGACCTGGGCTAGGAAAAAGATCTAATACCTACAGAGAGGCACATTTGA 683
OY 672 GACAAAGACATGTAGCCCAAGGAGGCTGGGCCAGAGAGACCACTCCAGCTCAACAG 731
DB 684 GACAAAGACATGTAGCCCAAGGAGGCTGGGCCAGAGAGACCACTCCAGCTCAACAG 743
OY 732 AACGTGAATCAACACCTGGGCCCAAGGCTAAGCTGAGGCCAAAGAGAGCATCTGCC 791
DB 744 AACGTGAATCAACACCTGGGCCCAAGGCTAAGCTGAGGCCAAAGAGAGCATCTGCC 803
OY 792 CGAGTGAATTCATCCCAAGAGACACCTGACACAGTCAACAGCAACAAACAGCAGTCC 851
DB 804 CGAGTGAATTCATCCCAAGAGACACCTGACACAGTCAACAGCAACAAACAGCAGTCC 863
OY 852 AGGAGCCCAATCCACAGAGGATCTTTGACAGAGCTGGGATTTCTAACAAACTCAGG 911
DB 864 AGGAGCCCAATCCACAGAGGATCTTTGACAGAGCTGGGATTTCTAACAAACTCAGG 922
OY 912 AACATCCATACAGCGATCAG-CCCTAGAGTGAAGGCAC-TTCTGATGGGTGCGAGA 969
DB 923 -ACTATCCCTACAGCGATCAGCCCTTAAGTGGAGGACCTTTCTGATGGGTGCGAGA 981
OY 970 TTCTGATGGGTGCGAGATGT-GGATCATCTTCAGACACAGATTTCTCAGCCAGATGG 1028
DB 982 TTCTGATGGGTGCGAGATGTGGGGATCTTCTCCGACCAGATTTCTCCAGCCCAATGG 1041
OY 1029 GTGGA---TGTTGGCCCAAGGAGTGAAGCTTCAAGG 1068
DB 1042 GGGGGATGGGGGGCCCAAGGATGAAGCTTAAGGCTTCAAGG 1084

```

RESULT 4
 B1826568 826 bp mRNA linear EST 04-OCT-2001
 LOCUS 603077016F1 NIH_MGC_119 Homo sapiens cdna clone IMAGE:516882 5'
 DEFINITION
 mRNA sequence.
 B1826568
 B1826568.1 GI:15938118
 KEYWORDS
 EST.
 SOURCE
 human.
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 1 (bases 1 to 826)
 NIH-MGC <http://mgc.nci.nih.gov/>.
 AUTHORS
 National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL
 Unpublished (1999)
 COMMENT
 Contact: Robert Strausberg, Ph.D.
 Email: cgapds-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cdna Library Preparation: Life Technologies, Inc.
 cdna Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: L14M1419 row: P column: 03
 High quality sequence stop: 823.

FEATURES

source

```

1.826
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:516882"
/clone_lib="NIH_MGC_119"
/tissue_type="medulla"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMV-Sport6; Site:1: NotI;
Site:2: EcoRV (destroyed); RNA source normal medulla from
anonymous male age 27. Library is oligo-dT primed and
directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.3 kb, insert size range
0.9-3 kb. Library is normalized and enriched for

```


OY	486	AAATCCAGAGGAAACCTGAGGCCCAAAACAAAATCTGCTTCAACAAGAGAAATTTCTGGCC	545
Db	421	AAATCCAGAGGAAACCTGAGGCCCAAAACAAAATCTGCTTCAACAAGAGAAATTTCTGGCC	480
OY	546	CACAGAGGCCCCACACAGCAATACCTTACATCCAAAGGGTGTGCTCACTCAACAG	605
Db	481	CACAGAGGCCCCACACAGCAATACCTTACATCCAAAGGGTGTGCTCACTCAACAG	540
OY	606	AAAGTCCTCCACAGAGGACCTGGGCTTGGAAAAGATCTATACTCAACAGAGCCAG	665
Db	541	AAAGTCCTCCACAGAGGACCTGGGCTTGGAAAAGATCTATACTCAACAGAGCCAG	600
OY	666	CATTGAGCAAAAGACATGTA--GCCAGCCAGGGCCCTGGGCCAGAGAGCCACTCCAGCT	724
Db	601	CATTGAGCAAAAGACATGTAAGGCCAGCCAGGGCCCTGGGCCAGAGAGCCACTCCAGCT	660
OY	725	CACAGAAGAGCTGAATCAACACCTCGGGCCAGGGCTAAACCTTGAAGCCAAAAGGGAGCCA	784
Db	661	CACAGAAGAGCTGAATCAACACCTCGGGCCAGGGCT--AACTGGAGCCAAAAGGGAGCCA	719
OY	785	TCTG--CCCCGATGATCTACATGCCAAGAGACCTGAAACGATGAGNAAACCAAAACAC	843
Db	720	TCTGCCCCGATGATCTACATGCCAAGAGACACTGAAACGATGAGNAAACCAAAACAC	779
OY	844	GCCAGTCCAGGAGGCCAAATCCAAACAGGG--ATCTTGAACAGAGCTGGGATTTTAACA	901
Db	780	GCCAGTCCAGGAGGCCAAATCCAAACAGGGATTTCTTTCAGAGAGCGTGGATTTCTTAC-	838
OY	902	AAACTTCAGGAATCTCCATACAGGATCAGCC	934
Db	839	AAACTTCAGGAATCTCCATACAGGATCAGCC	871

LOCUS	BI827706	905 bp	mRNA	linear	EST 04-OCT-2001
DEFINITION	60307612761 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5165959 5' ,				
ACCESSION	mRNA sequence.				
VERSION	BI827706				
KEYWORDS	BI827706.1	GI:15939256			
SOURCE	EST.				
ORGANISM	human.				
REFERENCE	Homo sapiens				
AUTHORS	Euhayyola; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
TITLE	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
JOURNAL	1 (bases 1 to 905)				
COMMENT	NIH-MGC http://mgc.nci.nih.gov/ .				
	National Institutes of Health, Mammalian Gene Collection (MGC)				
	Unpublished (1999)				
	Contact: Robert Strausberg, Ph.D.				
	Email: cgabbs@mail.nih.gov				
	Tissue Procurement: Life Technologies, Inc.				
	cDNA Library Preparation: Life Technologies, Inc.				
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)				
	DNA Sequencing by: Incyte Genomics, Inc.				
	Clone distribution: MGC clone distribution information can be				
	found through the I.M.A.G.E. Consortium/LNLN at:				
	http://image.llnl.gov				
	Plate: L14M1412 row: f column: 08				
	High quality sequence stop: 890.				
FEATURES	Location/Qualifiers				
source	1..905				

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5165959"
/clone_1id="N1H.MGC.119"
/tissue_type="medulla"
/lab_host="DH10B"
/notes="Organ: brain; Vector: PCMV-SpOET6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source normal medulla from anonymous male age 27. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.3 kb, insert size range

```

0.9-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 013. Note: this is a NIH-MGC Library."

Query Match	Similarity	96.7%	Score	78.4	DB	13	Length	905
Best Local	Similarity	96.7%	Pred	No. 2.4e-167				
Matches	876	Conservative	0	Mismatches	20	Indels	10	Gaps
QY	21	GGCACAATAGCCCCCTCCATATTCCTATCCCTCCACCAAGAAAGAGGCGACAGATTTCATT	80					
Db	1	GGCACAATAGCCCTCTCTATTTCTTCTTAATCCTCCACCAAGAAAGAGGCGACAGATTTCATT	60					
QY	81	ACTTGTGGGGGGCCAGCTGTGTATCGGCCCACTGCCAGCTGGCTTTAAAAAGAAAGCACT	140					
Db	61	ACTTGTGGGGGGCCAGCTGTGTATCGGCCCACTGCCAGCTGGCTTTAAAAAGAAAGCACT	120					
QY	141	GATGCTAGGATGGAATGAAACCCAGAGAGAAATGCCATCATCAGAGAATCAATAGAGATC	200					
Db	121	GATGCTAGGATGGAATGAAACCCAGAGAGAAATGCCATCATCAGAGAATCAATAGAGATC	180					
QY	201	TGTGAAGAGAGAGGCGCTGGGTGGAGCCCAAGAGATAGAACTGAGATCAATATCTC	260					
Db	181	TGTGAAGAGAGAGGCGCTGGGTGGAGCCCAAGAGATAGAACTGAGATCAATATCTC	240					
QY	261	CGTGAAGAGAAATTAACAATGAGACCAAGGTTCTTAAGTCAGTGTCTAGATGAGACTGGCAC	320					
Db	241	CGTGAAGAGAAATTAACAATGAGACCAAGGTTCTTAAGTCAGTGTCTAGATGAGACTGGCAC	300					
QY	321	CTGAACCAACACCAAGGCGCTTATACCCCGCTAGAGCCCTGGGCCAGAAAAACACCATAG	380					
Db	301	CTGAACCAACACCAAGGCGCTTATACCCCGCTAGAGCCCTGGGCCAGAAAAACACCATAG	360					
QY	381	CCGACCCGAATCGAAGAGCTGCGAGGGATCCCAATATCCCAAGAAAGCCGCTTCAAAAC	440					
Db	361	CCGACCCGAATCGAAGAGCTGCGAGGGATCCCAATATCCCAAGAAAGCCGCTTCAAAAC	420					
QY	441	AAAGACCCCTCACCCAGAGAGAGACCCCTGCACCAACATGATCTGTATCCCGAAGAGAAC	500					
Db	421	AAAGACCCCTCACCCAGAGAGAGACCCCTGCACCAACATGATCTGTATCCCGAAGAGAAC	480					
QY	501	CTAGAGCCCAACAAAATCTGCTTACAAGAGAAATTTCTTGCCCAACAGAACCCCGCAC	560					
Db	481	CTAGAGCCCAACAAAATCTGCTTACAAGAGAAATTTCTTGCCCAACAGAACCCCGCAC	540					
QY	561	CACACCAATCACCTTACATCATCAAGAGGTGCTCTCACTCAATCAAGAGAAAGTGGCTCCAGC	620					
Db	541	CACACCAATCACCTTACATCCAAAGGTGCTCTCACTCAATCAAGAGAAAGTGGCTCCAGC	600					
QY	621	AGGGAACCTGGGCTAGGAAAAAGATCTATTAATCAACAGAGAGCCAGCATTTGAGACAAAGAC	680					
Db	601	AGGGAACCTGGGCTAGGAAAAAGATCTATTAATCAACAGAGAGCCAGCATTTGAGACAAAGAC	660					
QY	681	ATGTAGCCCAAGCAGAGGCGCTGGGCCAGAGAGACCACTCACTCAACAAAGAAAGTGAAAT	740					
Db	661	ATGTAGCCCAAGCAGAGGCGCTGGGCCAGAGAGACCACTCACTCAACAAAGAAAGTGAAAT	718					
QY	741	CAACACCTGGGCGCCAGAGCTTAACCTGAGAGCAAAAAGGAGGCC--ATCTGGCCCGAGCTG	797					
Db	719	CAACACCTGGGCGCCAGAGCTTAACCTGAGAGCAAAAAGGAGGCCATCTGCTCCGAGCTTG	778					
QY	798	AATCTACATCCCA--GAGACACCTGAA--CACTGAGACAGAGAAACAGCCAGTCC--AG	853					
Db	779	AATCTACATCCCAAGAGAGACACCTGAAACAGTGTTCAGAAACAGATATCAACAGCCAGTCCAG	838					
QY	854	GGAGCAATATCCCAACAGAGAGATCTTGTGACAGAGCTGGGATTTCTTAACA--AAACTTCAGGA	912					
Db	839	GGAGCAATATCCCAACAGAGAGATCTTGTGACAGAGCTGGGATTTCTTAACAACAACTTCAGGA	898					
QY	913	ACTATC 918						

Db 899 ACTATC 904

RESULT 7
BI827559
LOCUS
DEFINITION BI827559 870 bp mRNA linear EST 04-OCT-2001
603073524F1 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:516578 5',
mRNA sequence.
ACCESSION BI827559
VERSION BI827559.1 GI:15939096
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 870)
NIH-MGC <http://mgc.nci.nih.gov/>,
National Institutes of Health, Mammalian Gene Collection (MGC)
UNPUBLISHED (1999)
CONTACT: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM1141 row: 0 column: 05
High quality sequence stop: 848.

FEATURES
source
1. 870
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="516578"
/clone_lib="NIH_MGC_119"
/tissue_type="medulla"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMV-SPORT6; Site: 1: NotI;
Site: 2: EcoRV (destroyed); RNA source normal medulla from
anonymous male age 27. Library is oligo-dT primed and
directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.3 kb, insert size range
0.9-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 013. Note:
this is a NIH MGC library."

BASE COUNT 287 a 234 c 221 g 128 t
ORIGIN

Query Match 20.6%; Score 782.6; DB 13; Length 870;
Best Local Similarity 97.8%; Pred. No. 4.9e-167;
Matches 846; Conservative 0; Mismatches 14; Indels 5; Gaps 5;

QY 55 CCAAGAAAGAGCAGCAGAGTTCATTAGTGGGGCCAGCTGTATGGCCACT-G 113
1 CCAAGAAAGAGCAGCAGAGTTCATTAGTGGGGCCAGCTGTATGGCCACTGG 60
Db 114 CCAAGTGCCTTAAAGAGAGCAGAGTGTAGATGGAGTGAACCCAGAGGAGT 173
61 GGGGCTGCCTTAAAGAGAGCAGAGTGTAGATGGAGTGAACCCAGAGGAGT 120
QY 174 GCCATCATGAGATCAATGAGATCTGTGAAGAGAGAGGGCTGGAGCCAGAA 233
121 GCCATCATGAGATCAATGAGATCTGTGAAGAGAGAGGGCTGGAGCCAGAA 180
QY 234 GGATGAGACCTGGAATCAATATCTCCGTGAGGGAATTAACATGAGCCAGTTCTA 293
181 GGATGAGACCTGGAATCAATATCTCCGTGAGGGAATTAACATGAGCCAGTTCTA 240
QY 294 AGTCAGTGTAGTCACTGCAACCTGAACACACAGAGCCCTATTAACCCCGTAG 353
241 AGTCAGTGTAGTCACTGCAACCTGAACACACAGAGCCCTATTAACCCCGTAG 300

QY 354 AGCTGGCCAGAAAAGACACCCATAGCCAGCCAGAACTCTGACAGGATCCA 413
Db 301 AGCTGGCCAGAAAAGACACCCATAGCCAGCCAGAACTCTGACAGGATCCA 360
QY 414 ATACCAGAGAAAGCTGCTTCAAAACCAAGACCCCTACCCAGAGAGACCCCTGCAC 473
Db 361 ATACCAGAGAAAGCTGCTTCAAAACCAAGACCCCTACCCAGAGAGACCCCTGCAC 420
QY 474 AACATGATGCTGAATCCCAAGAGAACCTAGAGCCCAAAATGCTTCAAGAGG 533
Db 421 AACATGATGCTGAATCCCAAGAGAACCTAGAGCCCAAAATGCTTCAAGAGG 480
QY 534 AATTTCTGGCCCAAGAAAGCCGACACAGCAATACCTTACATCAAAAGGCTGTC 593
Db 481 AATTTCTGGCCCAAGAAAGCCGACACAGCAATACCTTACATCAAAAGGCTGTC 540
QY 594 TCACTCAACAGAAAGCTGCTCCAGCAGGAGACCTGGGCTAGAGAAAGATCTATAAC-T 652
Db 541 TCACTCAACAGAAAGCTGCTCCAGCAGGAGACCTGGGCTAGAGAAAGATCTATAACGT 600
QY 653 CAACAGAGGACAGATGAGCAAAAGACATGTAGCCAGCCAGGACCTGGGCCAGAGAG 712
Db 601 CAACAGAGGACAGATGAGCAAAAGACATGTAGCCAGCCAGGACCTGGGCCAGAGAG 660
QY 713 CCAACTCCAGCTCAACAGAAAGCTGAATCAA-CACCTGCGCCAGGCTAAACCTGGAGC 771
Db 661 CCAACTCCAGCTCAACAGAAAGCTGAATCAAAGCAGCTGGGGCAGGCTAAACCTGGAGC 720
QY 772 CAAAAGGAGCCATCTGCCCCGACTGAATCTACATCCCAAGACACCTTAACAGTCA 831
Db 721 CAAAAGGAGCCATCTG-CGCCAGTGAATCTAGCT-CCAAAGACACCTGAACAGTCA 778
QY 832 CAAGCAAAAGAGCAGATCCAGAGAGCCAAATCAAGCAGAGATCTTGACAGAGCTGGG 891
Db 779 CAAGCAAAAGAGCAGATCCAGAGAGCCAAATCAAGCAGAGATCTTGACAGAGCTGGG 838
QY 892 ATTTCTAACAAACTTCAGAGACTA 916
Db 839 ATTTCTAACAAACTTCAGAGACTA 863

RESULT 8
BI561175 871 bp mRNA linear EST 05-SEP-2001
LOCUS BI561175
DEFINITION 603253712F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:5296155 5',
mRNA sequence.
ACCESSION BI561175
VERSION BI561175.1 GI:15448489
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 871)
NIH-MGC <http://mgc.nci.nih.gov/>,
National Institutes of Health, Mammalian Gene Collection (MGC)
UNPUBLISHED (1999)
CONTACT: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shihaki
Toshiyuki and Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM11749 row: 9 column: 04
High quality sequence stop: 820.

FEATURES
source
1. 871
Location/Qualifiers
/organism="Homo sapiens"

```

/db_xref="taxon:9606"
/clone="IMAGE:5296155"
/clone_lib="NIH_MGC_97"
/lab_host="DH10B"
/Note="Organ: testis; Vector: pBluescriptR (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTTNN-3',
size-selected for average insert size 2.2 kb and
normalized to 10^5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIMH/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT      281 a      242 c      215 g      133 t
ORIGIN

```

```

Query Match      20.1%; Score 764.8; DB 13; Length 871;
Best Local Similarity 96.6%; Pred. No. 5.5e-163;
Matches 835; Conservative 0; Mismatches 22; Indels 7; Gaps 5;

OY 1 CTTCTTTGTAAGAGAGTCTAGGACATAGCCCTCTCTATCTCTATCTCTCCACCAAG 60
    |||||||
DB 6 CTTCTTTGTAAGAGAGTCTAGGACATAGCCCTCTCTATCTCTATCTCTCCACCAAG 65
    |||||||

OY 61 AAAGAGGACACAGTTCATTACTTAGTGGGGCCACGCTGTGATCGGCCAACTGCCAGT 120
    |||||||
DB 66 AAAGAGGACACAGTTCATTACTTAGTGGGGCCACGCTGTGATCGGCCAACTGCCAGT 125
    |||||||

OY 121 CCTTAAAAAGAGACCACTGATGCTAGATGGAATGAACCCAGAGAGAAAGTCCATCA 180
    |||||||
DB 126 CCTTAAAAAGAGACCACTGATGCTAGATGGAATGAACCCAGAGAGAAAGTCCATCA 185
    |||||||

OY 181 TGAGGAATCAATGAGAGAGTCTGTAAGAGAGGCTGGGTGGAGCCCAAGAGATAGA 240
    |||||||
DB 186 TGAGGAATCAATGAGAGAGTCTGTAAGAGAGGCTGGGTGGAGCCCAAGAGATAGA 245
    |||||||

OY 241 ACCTGGAAGATCAATATCTCCGTGAGGGAATTAACAATGAGCCAGTTCCTAAGTCAGT 300
    |||||||
DB 246 ACCTGGAAGATCAATATCTCCGTGAGGGAATTAACAATGAGCCAGTTCCTAAGTCAGT 305
    |||||||

OY 301 GCTTAAAGTCAAGCTGGCAACCTGAACACACAGAGGCTTATTAACCCGCTAGAGGCTGG 360
    |||||||
DB 306 GCTTAAAGTCAAGCTGGCAACCTGAACACACAGAGGCTTATTAACCCGCTAGAGGCTGG 365
    |||||||

OY 361 GCCAAGAAAAGAACCCATATGAGCCGACGCAATGAGAGTCTGCAAGATCCATATCCCA 420
    |||||||
DB 366 GCCAAGAAAAGAACCCATATGAGCCGACGCAATGAGAGTCTGCAAGATCCATATCCCA 425
    |||||||

OY 421 ACAGAGAGCTGCTTAAACCAAGACCCCTCAACCCAGAGAGACCCCTGCACACATGA 480
    |||||||
DB 426 ACAGAGAGCTGCTTAAACCAAGACCCCTCAACCCAGAGAGACCCCTGCACACATGA 485
    |||||||

OY 481 TGCTGAATCCCGAAGAACTAGAGCCCAACAAAATCTGCTTCAAGAGAGAAATTTCT 540
    |||||||
DB 486 TGCTGAATCCCGAAGAACTAGAGCCCAACAAAATCTGCTTCAAGAGAGAAATTTCT 545
    |||||||

OY 541 TGCCCCAAGAGAGCCGACACCAATCAATCACTTATGATCCAAAGGCTGC-TGCTCACTC 599
    |||||||
DB 546 TGCCCCAAGAGAGCCGACACCAATCAATCACTTATGATCCAAAGGCTGC-TGCTCACTC 605
    |||||||

OY 600 AACAGAGAGTCTCTCCAGAGGAGCTGGGCTAGAGAAAGATCTATTAACCAACAG 659
    |||||||
DB 606 AACAGAGAGTCTCTCTCCAGAGGAGCTGGGCTAGAGAAAGATCTATTAACCAACAG 665
    |||||||

OY 660 AGCCAGCATTTGAGCAAAAGACATGTAGCCAGGCCAGGCTTGGGGCCAGAGAGAGCCA 725
    |||||||
DB 666 AGCCAGCATTTGAGCAAAAGACATGTAGCCAGGCCAGGCTTGGGGCCAGAGAGAGCCA 725
    |||||||

OY 717 CTCGAGCTCAACAAGAGTGAATCAACACCTGCGGCCAGGCTTAACCTGAGAGCCAAA 776
    |||||||
DB 726 CTCGAGCTCAACAAGAGTGAATCAACACCTGCGGCCAGGCTTAACCTGAGAGCC-AAA 784
    |||||||

OY 777 GGGAGCATCTCTGCCGACTGAATCTACATCCAGAGAGACACTGAACGTGAGACAGC 836
    |||||||

```

```

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||
DB 785 GGGAGCATCTGCCCCGACTG-ATCTAGTCTCCAGAGA-ACCTGACAGTCACGACGC 842
OY 837 AAACACGCCAGTCCAGGAGACCA 860
    ||||||| ||||||| |||
DB 843 AAACACGCCAGTCCAGGAGACCA 866

RESULT 9
BI830913
LOCUS
DEFINITION
603081040F1 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5172728 5',
mRNA sequence.
BI830913
ACCESSION
BI830913.1 GI:15942463
VERSION
BI830913.1
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 787)
NIH-MGC http://mgc.ncl.nih.gov/.
AUTHORS
National Institutes of Health, Mammalian Gene Collection (MGC).
TITLE
Unpublished (1999)
JOURNAL
Contact: Robert Strausberg, Ph.D.
COMMENT
Email: cgrabs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LMNL at:
http://image.lnl.gov
Plate: LMNL1429 row: P column: 09
High quality sequence stop: 784.
Location/Qualifiers
1.787
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5172728"
/clone_lib="NIH_MGC_119"
/tissue_type="medulla"
/lab_host="DH10B"
/Note="Organ: brain; Vector: pCMV-Sport6; Site_1: NotI;
Site_2: EcoRV (destroyed); RNA source normal medulla from
anonymous male age 27. Library is oligo-dT primed and
directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.3 kb, insert size range
0.9-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(invitrogen). Research Genetics tracking code 013. Note:
this is a NIH_MGC Library."
BASE COUNT      254 a      219 c      200 g      114 t
ORIGIN

Query Match      19.9%; Score 757; DB 13; Length 787;
Best Local Similarity 99.1%; Pred. No. 3.1e-161;
Matches 782; Conservative 0; Mismatches 5; Indels 2; Gaps 2;

OY 66 GGCACAGAGTTCATTACTTGTGGGGCCAGCTGTGATCGGCCCAACTGCGCTTGA 125
    |||||||
DB 1 GGCACAGAGTTCATTACTTGTGGGGCCAGCTGTGATCGGCCCAACTGCGCTTGA 60
    |||||||

OY 126 AAAAGGAAGACAGTGTAGTGAATGAACCCAGAGAGAGTGCATCATGAGG 185
    |||||||
DB 61 GAAAGGAAGACAGTGTAGTGAATGAACCCAGAGAGAGTGCATCATGAGG 120
    |||||||

OY 186 AATCAATGAGAGATCTGTGAAGAGAGGCTGGGTGGAGCCCAAGAGATAGAACTG 245
    |||||||
DB 121 AATCAATGAGAGATCTGTGAAGAGAGGCTGGGTGGAGCCCAAGAGATAGAACTG 180
    |||||||

OY 246 GAAGATCAATATCTCCGCTAGAGGAAATTAACAATGAGCCAGGTTCTAAGTCTCTA 305
    |||||||

```

Db	181	GAAAGATCAATATCTCTCCCGTAGGAGGAATTAACAATATGGAGCCAGGTTCTTAAGTCTGTCTA	240
Qy	306	GGTCAGACTGCGCAACCTGAACACACACAGAGGCGCTATTAACCCCGCTATGAGCGCTGGGCGAG	365
Db	241	GGTCAGACTGCGCAACCTGAACACACACAGAGGCGCTATTAACCCCGCTATGAGCGCTGGGCGAG	300
Qy	366	AAAACACACCCATTACCCAGCGCAGATTTGAAGACTCTTGCAGAGATTCGAATTAACCCACAGA	425
Db	301	AAAACACACCCATTACCCAGCGCAGATTTGAAGACTCTTGCAGAGATTCGAATTAACCCACAGA	360
Qy	426	AGCGCTCTTCAAAACCAAGACCCCTGACCCGACGAGGAGACCCCTGCACACATGATGCTG	485
Db	361	AGCGCTCTTCAAAACCAAGACCCCTGACCCGACGAGGAGACCCCTGCACACATGATGCTG	420
Qy	486	AATCCCAAGAGAACTTAGAGCCCAACAAAAATTGCTTCAACAAGAGAAATTTCTTGCC	545
Db	421	AATCCCAAGAGAACTTAGAGCCCAACAAAAATTGCTTCAACAAGAGAAATTTCTTGCC	480

Db	QY
719 CTGGCCCGACGTGAATCTACGTTCCAAAGAGACACCTGACAGCTCAGACAGCAAAACACGC	846 CAGTCCAGG 854
779 CAGTCCAGG 787	

[illegible]

```

FEATURES
source
location/Qualifiers
1. .788
/organism="Homo sapiens"

```

```

/db_xref="taxon:9606"
/clone="IMAGE:5163288"
/clone_lib="NIH_MGC_119"
/tissue_type="medulla"
/lab_host="DH10B"
/notes="Organ: brain; Vector: PCMV-SPORT6; Site:1; NotI:
Site:2; EcoRV (destroyed); RNA source normal medulla from
anonymous male age 27. Library is oligo-dT primed and
directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.3 kb, insert size range
0.9-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 013. Note
this is a NIH_MGC Library."
BASE COUNT      154 a      258 c      226 g      150 t
ORIGIN

```

Query Match	19.8%	Score 753.2;	DB 13;	Length 788;
Best Local Similarity	99.4%;	Pred. No. 2.3e-160;		
Matches 777; Conservative	0;	Mismatches 3;	Indels 2;	Gaps 2;

QY 1046 GGAGTGAAGCTAGCGCTCCAAAGGAAAATCTGGTTATAAGTGATGTCAGGATACAGTGGG 1105

07 **1106** ACGTGGCCACATGAGAAAAACCAGTGCTGGGAATTCAGACACACTACCAAGAATTACAGCCTTA **1165**

Db 63 ACGTCGCCACATGAGAAAAACCACTGCTCGAATCACACAGACACTACCAAGATACAGCCTCA 122

QY 1166 A-GGCTCATCCACACATGAGACTGGCGACCAATGACACAGTCGCTGGTACCTCGCGGA 1224

[illegible]

Db 183 GGACACATAGCCTTCTTCTCGAGCCA-GTCTCTGGGGAACGGCCAGCGGCTGTAGG 241

QY 1285 CGTATTGCCGGTGTACGGAGACAGGCGCTGGGGGCTGTGGAGCCGGGCCCTGGGCGGCTGCT 1344

1345 GGGTGTGGCGACCTCTTTGACCTGGACCCAGAGACACGGCCAAAGGGTACCGCAAGCTT 1404

Db 302 GGGTGTGGCGCACTCTTTGACCTGGACCGAGAGACACCGGCCAACGGGTACCGCAGCCT 361

1405 AGTACACACAGCCCCCTGCTGCCCTGCGACACTCTGTGACAAATCCGCTATGTGGCTC 1464

1465 CAACGCCGAGCATCTTCTTCGCGACGACCAACCTGGCCGAGCTGGAGGCGTACCT 1524

Db 422 CAACGGCGGACGATCTTCTTCGGCACGACCAACAACCTGGCGGAGCTGGAGGCTTACTT 481

1525 GGGCTGGCCCTCAACCCAGCCTCCGGCCTGTGCTACTACGCCCCAGCGCCTGTGTTACCA 1584
 1585
 1586
 1587
 1588
 1589
 1590
 1591
 1592
 1593
 1594
 1595
 1596
 1597
 1598
 1599
 1600
 1601
 1602
 1603
 1604
 1605
 1606
 1607
 1608
 1609
 1610
 1611
 1612
 1613
 1614
 1615
 1616
 1617
 1618
 1619
 1620
 1621
 1622
 1623
 1624
 1625
 1626
 1627
 1628
 1629
 1630
 1631
 1632
 1633
 1634
 1635
 1636
 1637
 1638
 1639
 1640
 1641
 1642
 1643
 1644
 1645
 1646
 1647
 1648
 1649
 1650
 1651
 1652
 1653
 1654
 1655
 1656
 1657
 1658
 1659
 1660
 1661
 1662
 1663
 1664
 1665
 1666
 1667
 1668
 1669
 1670
 1671
 1672
 1673
 1674
 1675
 1676
 1677
 1678
 1679
 1680
 1681
 1682
 1683
 1684
 1685
 1686
 1687
 1688
 1689
 1690
 1691
 1692
 1693
 1694
 1695
 1696
 1697
 1698
 1699
 1700
 1701
 1702
 1703
 1704
 1705
 1706
 1707
 1708
 1709
 1710
 1711
 1712
 1713
 1714
 1715
 1716
 1717
 1718
 1719
 1720
 1721
 1722
 1723
 1724
 1725
 1726
 1727
 1728
 1729
 1730
 1731
 1732
 1733
 1734
 1735
 1736
 1737
 1738
 1739
 1740
 1741
 1742
 1743
 1744
 1745
 1746
 1747
 1748
 1749
 1750
 1751
 1752
 1753
 1754
 1755
 1756
 1757
 1758
 1759
 1760
 1761
 1762
 1763
 1764
 1765
 1766
 1767
 1768
 1769
 1770
 1771
 1772
 1773
 1774
 1775
 1776
 1777
 1778
 1779
 1780
 1781
 1782
 1783
 1784
 1785
 1786
 1787
 1788
 1789
 1790
 1791
 1792
 1793
 1794
 1795
 1796
 1797
 1798
 1799
 1800
 1801
 1802
 1803
 1804
 1805
 1806
 1807
 1808
 1809
 1810
 1811
 1812
 1813
 1814
 1815
 1816
 1817
 1818
 1819
 1820
 1821
 1822
 1823
 1824
 1825
 1826
 1827
 1828
 1829
 1830
 1831
 1832
 1833
 1834
 1835
 1836
 1837
 1838
 1839
 1840
 1841
 1842
 1843
 1844
 1845
 1846
 1847
 1848
 1849
 1850
 1851
 1852
 1853
 1854
 1855
 1856
 1857
 1858
 1859
 1860
 1861
 1862
 1863
 1864
 1865
 1866
 1867
 1868
 1869
 1870
 1871
 1872
 1873
 1874
 1875
 1876
 1877
 1878
 1879
 1880
 1881
 1882
 1883
 1884
 1885
 1886
 1887
 1888
 1889
 1890
 1891
 1892
 1893
 1894
 1895
 1896
 1897
 1898
 1899
 1900
 1901
 1902
 1903
 1904
 1905
 1906
 1907
 1908
 1909
 1910
 1911
 1912
 1913
 1914
 1915
 1916
 1917
 1918
 1919
 1920
 1921
 1922
 1923
 1924
 1925
 1926
 1927
 1928
 1929
 1930
 1931
 1932
 1933
 1934
 1935
 1936
 1937
 1938
 1939
 1940
 1941
 1942
 1943
 1944
 1945
 1946
 1947
 1948
 1949
 1950
 1951
 1952
 1953
 1954
 1955
 1956
 1957
 1958
 1959
 1960
 1961
 1962
 1963
 1964
 1965
 1966
 1967
 1968
 1969
 1970
 1971
 1972
 1973
 1974
 1975
 1976
 1977
 1978
 1979
 1980
 1981
 1982
 1983
 1984
 1985
 1986
 1987
 1988
 1989
 1990
 1991
 1992
 1993
 1994
 1995
 1996
 1997
 1998
 1999
 2000
 2001
 2002
 2003
 2004
 2005
 2006
 2007
 2008
 2009
 2010
 2011
 2012
 2013
 2014
 2015
 2016
 2017
 2018
 2019
 2020
 2021
 2022
 2023
 2024
 2025
 2026
 2027
 2028
 2029
 2030
 2031
 2032
 2033
 2034

1585 TCGGCCCCGGGGACTCTTCTTTAGGGCGACGAGGGGCTACCGCGGACCTTCTCCGGGA 1644
OY

Db 542 TCGGCCCCGGGTA CTCTTC TTGAGGGCGACGAGGGGGCTCACCGCGACTTCTCTCCGGGA 601

1645 GTATTCACGCTGCATTAAGGAGTCTTATGCGCGCTGCGTTCAGTTCCAGTTCACGCC 1704

[illegible]

Db 662 TGGCATCCGGCCATTCTGCAGACATCTCCATTGGGCTGGTGTCTTCCTTCGGGAGCAGCTACTA 721

1765 CAACGGCAGACAGGCGCTCTAGTGTGGCCGCCAGCTCTCTTACACAGCGCGCTT 1824

QY 1825 TG 1826
Db 782 AG 783

RESULT 11
B1520160 773 bp mRNA linear EST 29-AUG-2001
LOCUS 603071405F1 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5163360 5',
DEFINITION mRNA sequence.
ACCESSION B1520160
VERSION B1520160.1 GI:15344952
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 773)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabds-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLM11405 row: j column: 01
High quality sequence stop: 769.

FEATURES
Source
Location/Qualifiers
1..773
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5163360"
/clone_1lb="NIH_MGC_119"
/tissue_type="medulla"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: EcoRV (destroyed); RNA source normal medulla from
anonymous male age 27. Library is oligo-dT primed and
directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.3 kb. Insert size range
0.9-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 013. Note:
this is a NIH_MGC Library."

BASE COUNT 152 a 254 c 220 g 147 t

Query Match 19.1%; Score 726; DB 13; Length 773;
Best Local Similarity 99.5%; Pred. No. 3.4e-154;
Matches 770; Conservative 0; Mismatches 0; Indels 4; Gaps 4;

QY 1043 CAGGAGTGAAGCTAGCTCAAGGAATCTGTTATTAAGTGTGACAGTACAGT 1102
Db 1 CAGGAGTGAAGCTAGCTCAAGGAATCTGTTATTAAGTGTGACAGTACAGT 59
QY 1103 GGGAGCTGCCACATGAGAAACAGTGTGCGAATCACAGACATACAGAGTACAGCC 1162
Db 60 GGGAGCTGCCACATGAGAAACAGTGTGCGAATCACAGACATACAGAGTACAGCC 119
QY 1163 TCAAGCTTCATCCACATGAGACCTGCGCACATGACACAGTGTGACTGTGGCG 1222
Db 120 TCAAGCTTCATCCACATGAGACCTGCGCACATGACACAGTGTGACTGTGGCG 179
QY 1223 GAGGCAACATATGCTTCTTCGAGCCAGGCTGCGGGGAAAGCCGAGGGCTGTCA 1282
Db 180 GAGGCAACATATGCTTCTTCGAGCCA-GGTCCTGGGGGAAAGCCGAGGGCTGTCA 238
QY 1283 GGCGTTTTGGCGGTATCGGAGCAGGGCGCTGGGGCTGAGCCGCGCCTGTGG 1342

Db 239 GCGCTTTTCCGGGTATACGGAGACAGCGCTGGGGCTGGAGCCGGCCCTGGCGCCTG 298
QY 1343 CTGGGTGTGGCGCACCTCTTTGACCTTGACCTGACCCAGAGACACCGCCACAGGGTACCAGAC 1402
Db 299 CTGGGTGTGGCGCACCTCTTTGACCTTGACCTGACCCAGAGACACCGCCACAGGGTACCAGAC 358
QY 1403 CTAGTGCACACAGCCCGCTGCTGGCGACACTCTGACCAATCCCGCTATGTGGCC 1462
Db 359 CTAGTGCACACAGCCCGCTGCTGGCGACACTCTGACCAATCCCGCTATGTGGCC 418
QY 1463 TCCAAACCGCCGACAGCATCTTCTCCGACACAGCCACCAACTGCGGAGCTGAGAGCTTAC 1522
Db 419 TCCAAACCGCCGACAGCATCTTCTCCGACACAGCCACCAACTGCGGAGCTGAGAGCTTAC 478
QY 1523 CTGGCTCCCTCACCCAGCTCCCGGCTGTGTCTACTAGCCGCCAGCGCTGTGTTAAC 1582
Db 479 CTGGCTCCCTCACCCAGCTCCCGGCTGTGTCTACTAGCCGCCAGCGCTGTGTTAAC 538
QY 1583 AATCGGCGGGGGTACTCTTGTAGGGGAGAGGGGGCTCACCGCGACTTCTCCGG 1642
Db 539 AATCGGCGGGGGTACTCTTGTAGGGGAGAGGGGGCTCACCGCGACTTCTCCGG 598
QY 1643 GAGTA-TGTACAGCTGCATTAAGGATGCTTATAGCGCTGCTGGGCTTCCAGTTTAC 1701
Db 599 GAGTATGTGATCAGCTGATTAAGGATGCTTATAGCGCTGCTGGGCTTCCAGTTTAC 658
QY 1702 GCCTGCCATCCGGGCATCTCTGACAGACATTCCTCATTTGGGCTGTGTGTCGGGGAGCA 1761
Db 659 GCCTGCCATCCGGGCATCTCTGACAGACATTCCTCATTTGGGCTGTGTGTCGGGGAGCA 718
QY 1762 CTCAAAAGCAAGCAG-ACAGCGCTCAGTGTGGCGGCGAGCTCTCTTACCA 1814
Db 719 CTCAAAAGCAAGCAGACAGCGCTCAGTGTGGCGGCGAGCTCTCTTACCA 772

RESULT 12
B1832284 781 bp mRNA linear EST 04-OCT-2001
LOCUS 603080125F1 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5171623 5',
DEFINITION mRNA sequence.
ACCESSION B1832284
VERSION B1832284.1 GI:15943834
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 781)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabds-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLM11427 row: b column: 08
High quality sequence stop: 781.

FEATURES
Source
Location/Qualifiers
1..781
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5171623"
/clone_1lb="NIH_MGC_119"
/tissue_type="medulla"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: EcoRV (destroyed); RNA source normal medulla from

anonymous male age 27. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.3 kb, insert size range 0.9-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics Tracking code 013. Note: this is a NIH-MGC Library."

BASE COUNT 250 a 216 c 195 g 120 t
ORIGIN

Query Match 18.7%; Score 711.2; DB 13; Length 781;

Best Local Similarity 97.2%; Pred. No. 7.9e-151;

Matches 756; Conservative 0; Mismatches 18; Indels 4; Gaps 3;

15 GTCTAGGCATAGATGCCCCCTCTATTCCTAATCCTCCACCAA-AGAAAGAGCAGACA 73
|||||
Db GTCTAGGCATAGATGCCCCCTCTATTCCTAATCCTCCACCAAAGAGCAGACA 65
74 GTTCATTACTTACTTACTGGGGCCAGCTGTGATGGCCAACTGCCAGTCCCTTAAAGAA 133
|||||
Db GTTCATTACTTACTTACTGGGGCCAGCTGTGATGGCCAACTGCCAGTCCCTTAAAGAA 125
134 GACCAATGATGCTAGATGAGTGAACCCAAAGAGAGTGCATCATGAGGAATCAATG 193
|||||
Db GACCAATGATGCTAGATGAGTGAACCCAAAGAGAGTGCATCATGAGGAATCAATG 185
126 GACCAATGATGCTAGATGAGTGAACCCAAAGAGAGTGCATCATGAGGAATCAATG 195
194 AGAGATCTGTGAAGAGAGAGGGCTGGAGCCCAAGAGATAGAACTGGAGATCA 253
|||||
Db AGAGATCTGTGAAGAGAGAGGGCTGGAGCCCAAGAGATAGAACTGGAGATCA 245
186 AGAGATCTGTGAAGAGAGAGGGCTGGAGCCCAAGAGATAGAACTGGAGATCA 245
254 ATATCTCCGTGAGGGAATTAACAATGAGCCAGGTTCTTAAGTCACTGTCTAGTCAGAC 313
|||||
Db ATATCTCCGTGAGGGAATTAACAATGAGCCAGGTTCTTAAGTCACTGTCTAGTCAGAC 305
246 ATATCTCCGTGAGGGAATTAACAATGAGCCAGGTTCTTAAGTCACTGTCTAGTCAGAC 305
314 TGGCACTGTAACACACAGAGAGGCTATPACCCGCTGAGCTGGGCCAGAAAGACA 373
|||||
Db TGGCACTGTAACACACAGAGAGGCTATPACCCGCTGAGCTGGGCCAGAAAGACA 365
306 TGGCACTGTAACACACAGAGAGGCTATPACCCGCTGAGCTGGGCCAGAAAGACA 365
374 CCCATAGCCAGCCAGATGGAAGACTGTGACAGGATTCATACCAACAGAGAGCTGCT 433
|||||
Db CCCATAGCCAGCCAGATGGAAGACTGTGACAGGATTCATACCAACAGAGAGCTGCT 425
366 CCCATAGCCAGCCAGATGGAAGACTGTGACAGGATTCATACCAACAGAGAGCTGCT 425
434 TCACCAACAAAGACCCCTCACCCAGAGAGAGCCCTGACACATGATGCTGAATCCAG 493
|||||
Db TCACCAACAAAGACCCCTCACCCAGAGAGAGCCCTGACACATGATGCTGAATCCAG 485
426 TCACCAACAAAGACCCCTCACCCAGAGAGAGCCCTGACACATGATGCTGAATCCAG 485
494 AAGGAACCTAGAGCCCAACAAAATCTGCTCAAGAGAGAAATTTCTGCCCAACAGAG 553
|||||
Db AAGGAACCTAGAGCCCAACAAAATCTGCTCAAGAGAGAAATTTCTGCCCAACAGAG 545
486 AAGGAACCTAGAGCCCAACAAAATCTGCTCAAGAGAGAAATTTCTGCCCAACAGAG 545
554 CCCGACACAGAGATCACCCTTACATCAAAAGGCTGCTCACTCAACAGAGAGCTGCC 613
|||||
Db CCCGACACAGAGATCACCCTTACATCAAAAGGCTGCTCACTCAACAGAGAGCTGCC 604
546 CCCGACACAGAGATCACCCTTACATCAAAAGGCTGCTCACTCAACAGAGAGCTGCC 604
614 TCCACACAGAGAGCTGGGCTAGAGAAAAGATCTATACTCAACAGAGAGCAGATTGAGA 673
|||||
Db TCCACACAGAGAGCTGGGCTAGAGAAAAGATCTATACTCAACAGAGAGCAGATTGAGA 663
605 -TCCACACAGAGAGCTGGGCTAGAGAAAAGATCTATACTCAACAGAGAGCAGATTGAGA 663
674 CAAAGACATGTAGCCAGCCAGGAGGCTGGGCCA-GGAGAGCCACTCCAGCTCAACAGAG 732
|||||
Db CAAAGACATGTAGCCAGCCAGGAGGCTGGGCCAAGAGAGCCACTCCAGCTCAACAGAG 723
664 CAAAGACATGTAGCCAGCCAGGAGGCTGGGCCAAGAGAGCCACTCCAGCTCAACAGAG 723
733 AGCTGATCAACACAGCTGGGCGCCAGGCTAAACCTGAGAGCCAAAGAGAGCAGCTGCC 790
|||||
Db AGCTGATCAACACAGCTGGGCGCCAGGCTAAACCTGAGAGCCAAAGAGAGCAGCTGCC 781
724 AGCTGATCAACACAGCTGGGCGCCAGGCTAAACCTGAGAGCCAAAGAGAGCAGCTGCC 781

RESULT 13
AI337272

LOCUS AI337272 767 bp mRNA linear EST 18-MAR-1999
DEFINITION tb96h07.x1 NCI_CGAP Col6 Homo sapiens cDNA clone IMAGE:2062227 3'

similar to gb:L11706 cds1 HORMONE SENSITIVE LIPASE (HUMAN);, mRNA
sequence.

ACCESSION AI337272
VERSION AI337272.1 GI:4074199
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
TITLE NCI-CGAP http://www.ncbi.nlm.nih.gov/nclogap.
JOURNAL National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)

COMMENT
CONTACT: Robert Strausberg, Ph.D.
Email: cga@bbs-r@mail.nih.gov
Tissue Procurement: Ilan Kirsch, M.D., Michael R. Emmert-Buck, M.D.,
Ph.D.

CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/bbrp/image/image.html
Insert length: 3032 Std Error: 0.00
Seq primer: -40bp from Gibco
High quality sequence stop: 502.

FEATURES

source

1..767

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:2062237"

/clone_1lb="NCI_CGAP_Col6"

/tissue_type="colon tumor, RER+"

/lab_host="DH10B"

/note="Organ: colon; Vector: pT73D-Pac (Pharmacia) with a
modified polylinker; Site:1: Not I; Site:2: Eco RI;
Plasmid DNA from the normalized library NCI_CGAP_Col6 was
prepared, and ss circles were made in vitro. Following HAP
purification, this DNA was used as tracer in a subtractive
hybridization reaction. The driver was PCR-amplified cDNAs
from a pool of 5,000 clones made from the same library
(clonoids 1057416-1061255, and 1144584-1145351).
Subtraction by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 134 a 273 c 217 g 138 t 5 others
ORIGIN

Query Match 18.3%; Score 697; DB 9; Length 767;

Best Local Similarity 96.4%; Pred. No. 1.3e-147;

Matches 742; Conservative 0; Mismatches 25; Indels 3; Gaps 3;

1815 GCGGCGCTTTGGCATGACCCCGAGCTCGTGGGCTGAGTTTGAAGCGATCACACAGA 1874
|||||
Db 1 GCGGCGCTTTGGCATGACCCCGAGCTCGTGGGCTGAGTTTGAAGCGATCACACAGA 60
1875 ACCTGACGTGCACTCTTGAAAAGCTTCTGGAACATCACCGAGATGGAAGTCTATCGT 1934
|||||
Db ACCTGACGTGCACTCTTGAAAAGCTTCTGGAACATCACCGAGATGGAAGTCTATCGT 120
61 ACCTGACGTGCACTCTTGAAAAGCTTCTGGAACATCACCGAGATGGAAGTCTATCGT 120
1935 CTCTGGCAACATGGCATGGGCGACCGATGAGGTTAAGCGGCTGCTCAGCTGCCACCGG 1994
|||||
Db CTCTGGCAACATGGCATGGGCGACCGATGAGGTTAAGCGGCTGCTCAGCTGCCACCGG 180
121 CTCTGGCAACATGGCATGGGCGACCGATGAGGTTAAGCGGCTGCTCAGCTGCCACCGG 180
1995 AAGCCTTTGAGATGCCATGACTGCGACCCGACGCTCAAGGCTCAACATCTCAACCCCCAC 2054
|||||
Db AAGCCTTTGAGATGCCATGACTGCGACCCGACGCTCAAGGCTCAACATCTCAACCCCCAC 240
181 AAGCCTTTGAGATGCCATGACTGCGACCCGACGCTCAAGGCTCAACATCTCAACCCCCAC 240
2055 TGGCCACACAGAGCCCTGGGCGCGTCTGTGCAAGCTCATCTCTATGACCTGCGTGAAG 2114
|||||
Db TGGCCACACAGAGCCCTGGGCGCGTCTGTGCAAGCTCATCTCTATGACCTGCGTGAAG 300
241 TGGCCACACAGAGCCCTGGGCGCGTCTGTGCAAGCTCATCTCTATGACCTGCGTGAAG 300
2115 GACAGGACGTGAGGAGCTCAGAGCGCTGATTAAGTCCAAAGGCCCAAGGAGAGCCTGGAGC 2174
|||||
Db GACAGGACGTGAGGAGCTCAGAGCGCTGATTAAGTCCAAAGGCCCAAGGAGAGCCTGGAGC 360

OY	2175	TTGAGCGCGGCCCCAGCAGGACCGCGCGGTCCTGATATGACATTCCACGGGG	2234
Db	361	TGTGGCCGCGCCCCAGCAGGACCCCGCTCGCGGTCCTGATATGACATTCCACGGGG	420
OY	2235	GTGGCTTTGTGGCCAGACCTCCAGATCCACAGAGCCCTACTCTCAAGACTGGGCCAG	2294
Db	421	GTGGCTTTGTGGCCAGACCTTACAGATCCACAGAGCCCTACTCTCAAGACTGGGCCAG	480
OY	2295	AGCTGGGGCCCCCATCATCTCCATCGACTACTCCCTGGCCCCGTGAGGCCCTTCCCC	2354
Db	481	AGCTGGGGCCCCCATCATCTCCATCGACTACTCCCTGGCCCCGTGAGGCCCTT-CCCC	539
OY	2355	GTGGCGTGGAGAGTGCCTTCTCGCCTACTGTGCGGGCCATCAAGCACTGCCCTCTCTTG	2414
Db	540	GTGGCGTGGAGAGTGCCTTCTCGCCTACTGTGCGGGCCATCAAGCACTGCCCTCTCTTG	599
OY	2415	GCTCAACAGGGAGACGAATCTGCTTGGCGGGGAGAGTTCAGACGCGGGAACCTCTCTTCA	2474
Db	600	GCTCAAC- GGGGAACGAATCTGCTTGGC- GGGGACAGTTCAGACGCGGGAACCTCTCTTCA	657
OY	2475	CCGTGGCTTTCGGGCGACGACCTTACGGGGGTGGGGTGCAGATGGCATCATGGACCT	2534
Db	658	NCGTGGCTTTCGGGCGACGACCTTACGNNGTGGCGGTGCAGATGGCATTATGGCAGGAC	717
OY	2535	ACCGCGGCACATGCTGCAGCCTTGCACGCTCTCTCCCTCCGCGCTGCTGAGC	2584
Db	718	TACCGCGGCACATGCTGCAGCCTTGCACGCTCTCTCTCCCTCCGCGCTGCTGAGC	767

BC029301	BC029301	1367 bp	mRNA	linear	HTC 01-MAY-2002
LOCUS					
DEFINITION					
Accession					
BC029301					
LOCUS					
DEFINITION					
Accession					
BC029301					
LOCUS					
DEFINITION					
Accession					
BC029301					
LOCUS					
DEFINITION					
Accession					
BC029301					
LOCUS					
DEFINITION					
Accession					
BC029301					
LOCUS					
DEFINITION					
Accession					
BC029301					
LOCUS					
DEFINITION					
Accession					
BC029301					
LOCUS					
DEFINITION					
Accession					
BC029301					
LOCUS					
DEFINITION					
Accession					
BC029301					
LOCUS					
DEFINITION					
Accession					
BC029301					
LOCUS					
DEFINITION					
Accession					
BC029301					
LOCUS					
DEFINITION					
Accession					
BC029301					
LOCUS					
DEFINITION					
Accession					
BC029301					
LOCUS					
DEFINITION					
Accession					
BC029301					
LOCUS					
DEFINITION					
Accession					
BC029301					
LOCUS					
DEFINITION					
Accession					
BC029301					
LOCUS					
DEFINITION					
Accession					
BC029301					
LOCUS					
DEFINITION					
Accession					
BC029301					
LOCUS					
DEFINITION					
Accession					
BC029301					
LOCUS					
DEFINITION					
Accession					
BC029301					
LOCUS					
DEFINITION					
Accession					
BC029301					
LOCUS					
DEFINITION					
Accession					
BC029301					
LOCUS					
DEFINITION					
Accession					
BC029301					
LOCUS					
DEFINITION					
Accession					
BC029301					
LOCUS					
DEFINITION					
Accession					
BC029301					
LOCUS					
DEFINITION					
Accession					
BC029301					
LOCUS					
DEFINITION					
Accession					
BC029301					
LOCUS					
DEFINITION					
Accession					
BC029301					
LOCUS					
DEFINITION					
Accession					
BC029301					
LOCUS					

VERSION	BC029301.1	GI:20381052
KEYWORDS	HTC.	
SOURCE	Homo sapiens.	
ORGANISM	Homo sapiens	

REFERENCE
AUTHORS
TITLE
JOURNAL

1 (bases 1 to 1367)
Strausberg, R.
Direct Submission
Submitted (01-MAY-2002) National Institutes of Health, Mammalian

REMARK	COMMENT
NIH-MGC Project URL: http://mgc.nci.nih.gov Contact: MGC help desk	<p> T. J. Slamon, M.D., F.R.C.P. Director, Cancer Biomarkers Center, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA </p>

Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILLUM)
DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: amg@bcm.tmc.edu
Gunarane, P. H., Garcia, A. M., Lu, X., Huylk, S. W., Hale, S. M., Yoon, V. S., Kowis, C. R., Lawrence, S., Martin, R. G., Muzny, D. M., Richards, S., Gibbs, R. A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAK Plate: 51 Row: c Column: 24
 This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis
 This clone has the following problem: no polyA-tail.

FEATURES	location/Qualifiers
source	1. 1367
	/organism="Homo sapiens"
	/db_xref="taxon:9606"

```

/clone="IMAGE:5169931"
/tissue_type="Brain, adult medulla"
/clone_lib="NTH_MGC_119"
/lab_host="DH10B"
/note="vector: pCMV-SPORT6"
BASE COUNT      240 a      445 c      453 g      229 t
ORIGIN

```

Query Match	17.6%	Score 668.8	DB 11	Length 1367
Best Local Similarity	99.0%	Pred. No. 4.5e-141		
Matches 673	Conservative 0	Mismatches 7	Indels 0	Gaps 0

QY	1143	688	Db
	GACACTACACAGATACAGCCTCAAGGCTCATCCACAATGACCTGGGCACATGACAC	1202	
	GAAGAACAAGGCGCAAAACCTCAAGGCTCATCCACAACATGACCTGGGCACATGACAC	747	

QY	1203	AGTGGCTGTGACTCTGCGCGAGAGACAACATAGCCTTCTTCTGAGCCAGGCTCTGGGG	1262
Db	748	AGTGGCTGTGACTCTGCGCGAGAGACAACATAGCCTTCTTCTGAGCCAGGCTCTGGGG	807

QY	1263	AAACGGCCAGCGGCTGTCAGCGGCTTTTTCGCGGTGTCACGGGACACAGCGCGCTGGGGGCTCG	1322
Db	808	AAACGGCCAGCGGCTGTCAGCGGCTTTTTCGCGGTGTCACGGGACACAGCGCGCTGGGGGCTCG	867

QY	1323	AGCGGGCCCTGSGGCGCCCTGCTGGGTGTGGCGCACCTTTTGACCTGACCCAGAGACAC	1382
Db	868	AGCGGGCCCTGGGCGCCCTGCTGGGTGTGGCGCACCTTTTGACCTGAGCCCGAGAGAC	927

QY	CGGCGAACGGGTACCGCGACGCTAGTGCACACAGCCCGCTGCTGCGCGACCTCTCGGC	1442
1383	CGGCGAACGGGTACCGCGACGCTAGTGCACACAGCCCGCTGCTGCGCGACCTCTCGGC	
QY	CGGCGAACGGGTACCGCGACGCTAGTGCACACAGCCCGCTGCTGCGCGACCTCTCGGC	1442
Db	CGGCGAACGGGTACCGCGACGCTAGTGCACACAGCCCGCTGCTGCGCGACCTCTCGGC	987
928	CGGCGAACGGGTACCGCGACGCTAGTGCACACAGCCCGCTGCTGCGCGACCTCTCGGC	987

QY 1443 ACAAATCCCGCTATGTGGCCCTCCAAACCGCGCAGCATCTTCTTCGCGCAACGACCAACCC 1502

Db 988 ACAAATCCCGCTATGTGGCCCTCCAAACCGCGCAGCATCTTCTTCGCGCAACGACCAACCC 1047

Qy	1503	TGGCCGAGCTGGAGGGCTACCTGGCTGCCCTCACCAGCTCCGCTGTGTACTACG	1562
Db	1048	TGGCCGAGCTGGAGGGCTACCTGGCTGCCCTCACCAGCTCCGCTGTGTACTACG	1107

Qy	1563	CCAGCGCGCTGTTGGTATACCAATCGGCCGGGGGTACTCTTCTTTGAGGGCGAGCGAGGGGC	1622
Dp	1108	CCCCGCGCTGCTGGTATTACCAATCGGCCGGGGGGTACTCTTCTTTGAGGGCGAGCGAGGGGC	1167

QY	1623	TCACCGCGGACTCTCTCCGGAGATATGCACGCTGCATAAAGGATGCTCTATATGGCCGCT	1682
DB	1168	TCACCGCGGACTCTCTCCGGAGATATGCACGCTGCATAAAGGATGCTCTATATGGCCGCT	1227

0Y	GCCTGGGCTTCAGTTCACGCGCTGCATTCCTCTCGAGACCATCTCCATTGGGC	1742
1683	GCCTGGGCTTCAGTTCACGCGCTGCATTCCTCTCGAGACCATCTCCATTGGGC	1742
1228	GCCTGGGCTTCAGTTCACGCGCTGCATTCCTCTCGAGACCATCTCCATTGGGC	1287

0y 1743 TGTGTCTCTTGGGGAGCAGTACAAACGCAAGCAGACAGCGCTCAGTGTGGCGCCAGCT 1802
1288 TTTTCTCTCTTGGGGAGCAGTACAAACGCAAGCAGACAGCGCTCAGTGTGGCGCCAGCT 1347

QY	CTCTCTTCACCAAGCGCCG	1822
Ob	CTCTCTTCACCAAGCGCCG	1267

RESULT 15

LOCUS	B1826036	663 bp	mRNA	linear	EST 04-OCT-2001
DEFINITION	603076441.1	NIH_MGC_119	Homo sapiens	CDNA	clone IMAGE:5168147 5'
REMARKS	mRNA sequence.				

```

VERSION      BI826036.1  GI:15937586
KEYWORDS
SOURCE       human.

```

Eukaryota; Metazoa; Chordata; Craniata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

REFERENCE 1 (bases 1 to 663)
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabs-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
 Clone distribution: MGC clone distribution information can be
<http://image.llnl.gov>
 found through the I.M.A.G.E. Consortium/LLNL at:
 Plate: L14M1418 row: a column: 12
 High quality sequence stop: 661.
 Location/Qualifiers

FEATURES
 source 1..663
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5168147"
 /clone_1ib="NIH-MGC_119"
 /tissue_type="medulla"
 /lab_host="DH10B"
 /note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
 Site_2: EcoRV (destroyed); RNA source normal medulla from
 anonymous male age 27. Library is oligo-dT primed and
 directionally cloned (EcoRV site is destroyed upon
 cloning). Average insert size 1.3 kb. Insert size range
 0.9-3 kb. Library is normalized and enriched for
 full-length clones and was constructed by C. Gruber
 (Invitrogen). Research Genetics tracking code 013. Note:
 this is a NIH-MGC Library."

BASE COUNT 216 a 179 c 169 g 99 t
 ORIGIN

Query Match 17.1%; Score 652; DB 13; Length 663;
 Best Local Similarity 99.8%; Pred. No. 2,1e-137;
 Matches 663; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

```

QY 57 AAGAGAAAGGACACAGATTCTTACTTACTGAGGGCCAGCTGTGATGAGCCAACTGCCA 116
   |||||||
Db 1 AAGAGAAAGGACACAGATTCTTACTTACTGAGGGCCAGCTGTGATGAGCCAACTGCCA 60
QY 117 GCTGCTTAAAGAGAACAGCAGTGTAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 176
   |||||||
Db 61 GCTGCTTAAAGAGAACAGCAGTGTAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 120
QY 177 ATCATGAGAGATCAATGAGATCTGTGAAGAGAGAGGGCTGGTGGAGAGCCAGAAAGA 236
   |||||||
Db 121 ATCATGAGAGATCAATGAGATCTGTGAAGAGAGAGGGCTGGTGGAGAGCCAGAAAGA 180
QY 237 TAGAAGCTGGAAGATCAATATCTCCGTGAGGGAATATACATGAGCCAGGTTCTAAGT 296
   |||||||
Db 181 TAGAAGCTGGAAGATCAATATCTCCGTGAGGGAATATACATGAGCCAGGTTCTAAGT 240
QY 297 CAGTGTAGGCTGAGCTGCAACCTGACACACACAGAGGCTATTAACCCGCTAGAGC 356
   |||||||
Db 241 CAGTGTAGGCTGAGCTGCAACCTGACACACACAGAGGCTATTAACCCGCTAGAGC 300
QY 357 CTGGGCCAGAAAAGACACCCATAGCCAGCAGCAATGAAAGCTTGCAGAGATCCAAATA 416
   |||||||
Db 301 CTGGGCCAGAAAAGACACCCATAGCCAGCAGCAATGAAAGCTTGCAGAGATCCAAATA 360
QY 417 CCCAAGAGAGCTGCTTAAACCAAGAGCCCTCAACCCAGAGAGAGACCCCTGCACAAAC 476
   |||||||
Db 361 CCCAAGAGAGCTGCTTAAACCAAGAGCCCTCAACCCAGAGAGAGACCCCTGCACAAAC 420
QY 477 ATGATGCTGATCCCGAAGAGAACTTAGAGCCACAAAATCTGCTCACAAGAGGAAT 536
   |||||||
Db 421 ATGATGCTGATCCCGAAGAGAACTTAGAGCCACAAAATCTGCTCACAAGAGGAAT 480
QY 537 TTCTTGCCCAAGAGCCGACCCAGACCAATCACTTACATCCAAAGGCTGCTGCTCA 596
   |||||||

```

```

Db 481 TTCTTGCCCAAGAGCCGACCCAGACCAAGCAATCACTTACATCCAAAGGCTGCTGCTCA 540
QY 597 CTCACAGAGAGAGCTGCTCCCTCCAGCAGAGACCTGGGCTAGAGAAAGATCTATTAAGTCAAC 656
   |||||||
Db 541 CTCACAGAGAGAGCTGCTCCCTCCAGCAGAGACCTGGGCTAGAGAAAGATCTATTAAGTCAAC 600
QY 657 AGGAGCCAGCAATTTAGACAAAGACATGTAGCCAGCAGAGGCTGGCCAGAGAGAGCCAC 716
   |||||||
Db 601 AGGAGCCAGCAATTTAGACAAAGACATGTAGCCAGCAGAGGCTGGCCAGAGAGAGCCAC 659
QY 717 CTGC 720
   ||||
Db 660 CTGC 663

```

Search completed: January 3, 2003, 07:06:07
 Job time : 3339 secs